Land-Surface Modelling and Climate Simulations: Results over the Australian Region from sixteen AMIP2 models

Huqiang Zhang[1], A. Henderson-Sellers[2], P. Irannejad[2], S. Sharmeen[2], T. Phillips[3], K. McGuffie[4]

[1]Bureau of Meteorology Research Centre, Melbourne, Australia

[2]ANSTO Environment, Sydney, Australia

[3]PCMDI, Livermore, CA, USA

[4] Applied Physics, University of Technology, Sydney, Australia

(10 April, 2002)

(AMIP2 subproject-12 internal report)

Corresponding Author:

Dr. Huqiang Zhang

BMRC

GPO Box 1289K

VIC 3001, Australia

Email: h.zhang@bom.gov.au

Abstract

This report presents analyses of sixteen models from the Atmospheric Model Intercomparison Project II (AMIP2) over the Australian region. It is focused on assessing how well surface climate and fluxes over this region are simulated in current Atmospheric General Circulation Models (AGCMs) forced by observed sea surface temperatures (SSTs). The importance of land-surface modelling on model predictability is also investigated. In this preliminary analysis, the Bureau of Meteorology (BoM) observational rainfall, temperature and surface evapotranspiration datasets are used in validating surface climatologies simulated by the 16 models. Specifically, the Linear Error in Probability Space (LEPS) score is calculated in assessing the skill of the models in simulating surface climate anomalies for the 17-year period (1979 to 1995). Numerous model differences are seen with some aspects of the model performance being linked to the complexity of land-surface schemes used. The connection between model skill in simulating surface climate anomalies and surface flux anomalies is explored.

Lag-correlation analysis is conducted. Results reveal that "climatic memory" derived from land-surface processes (e.g. soil moisture) has different features in the sixteen models: some models show rapid feedback processes between land-surface and the overlying atmosphere, while others show slowly varying processes in which anomalous surface conditions have impacts on the model integrations on longer time-scales. It is found that models with simple bucket-type scheme tend to have a more rapid decay rate in the retention of soil moisture anomalies, and therefore, soil moisture conditions have a much weaker influence on forecasting surface climate anomalies. This study suggests that land-surface modelling has the potential to influence AGCM predictability on seasonal and even longer time scales.

1. Introduction

Over the last decade, the Atmospheric Model Intercomparison Project (AMIP) has evolved as a standard framework for evaluating state-of-art Atmospheric General Circulation Models (AGCMs) in their simulations of current climate system. Gates et al. (1999) summarised the scientific findings and contributions from the first phase of AMIP. Comparing model simulations with observations and contrasting results among models have helped to identify the strengths and weaknesses in current AGCMs and thus pinpoint areas where efforts are needed to improve the ability to simulate the current climate and predict future climates. AMIP1 outcomes are largely achieved by the analyses of model dynamical and physical processes through diagnostic subprojects (Gates et al., 1999). The AMIP I diagnostic subproject 12, forms part of the Project for Intercomparison of Land-surface Parameterization Schemes (PILPS) (Henderson-Sellers et al., 1996), is dedicated to evaluating AGCM simulations of surface energy and water budgets and their components, and to assessing the role of land-surface parameterisations in AGCM simulations. As summarized by Henderson-Sellers (1999), the major findings from AMIP I subproject 12 were that: (i) no "best" land-surface simulation could be identified with every model showing unsatisfactory results in some respect (Love and Henderson-Sellers, 1994); (ii) some models failed to conserve surface energy and water balances, with pronounced trends in moisture stores (Love et al., 1995); and (iii) energy and water partitioning showed larger model discrepancies than those seen in PILPS offline intercomparisons (Irannejad et al., 1995; Qu and Henderson-Sellers, 1998).

Assessing the role of land-surface parameterisations in AGCM simulations remains a key area in weather and climate modelling (e.g., Zhang et al., 2001; 2002) and is further pursued in the AMIP

phase II (AMIP2) experiments. Phillips et al. (2002) outlined the overall scientific plans in AMIP2 subproject 12. Like its predecessor, this is aimed to assess the degree to which model performance in simulating land-surface climate is related to the complexity of land-surface schemes employed. In contrast with AMIP I, there are a number of advantages of analysing AMIP2 experiments in relation to the study of land-surface processes in global climate system. These include: (i) a greater variety of complexity in land-surface schemes employed; (ii) better control of the model initialisation and spin-up processes; and (iii) more variable outputs which help to characterize land-surface processes. In a pilot study of this AMIP2 diagnostic subproject, Phillips et al. (2000) proposed some scientific approaches for the exploration of spatio-temporal variability of land-surface simulations from AMIP2.

This study is focused on assessing the performance of AMIP2 models over the Australian region (see e.g. Figure 2). There are two main goals: (a) to deliver an evaluation of current AGCMs' simulations of observed climate over this region; and (b) to try to establish whether, and potentially, how, land-surface processes and parameterisations affect the model predictability of climate anomalies on seasonal and longer time scales (e.g., Zhang and Frederiksen, 2001). As a first step, the focus of the preliminary analysis is centred on three questions (i) how different the models are in simulating key surface climate variables such as precipitation and surface temperature; (ii) how different the models are in simulating land-surface fluxes (latent and sensible heat fluxes); and (iii) whether there is any connection between the model skill in simulating surface flux anomalies and in predicting surface climate anomalies.

There are sound scientific reasons for proposing that land-surface modelling could affect model predictability. Primarily, different representation of land-surface processes may affect the time scale of feedbacks between the land-surface and the atmosphere. For example, considering how soil moisture is simulated in GCMs, three main characterisations can be identified. In a one-layer bucket-type model (e.g., Manabe, 1969), soil moisture is governed by

$$rH \frac{dw}{dt} = P - E - R - D \tag{1}$$

where H is soil depth, Δ is water density, w is volumetric soil water content, P is precipitation, E is surface evaporation, R is surface runoff and D is drainage term. In this type of parameterisation, the occurrence of runoff is determined by whether soil moisture is above a critical value (field capacity). No runoff is simulated until the whole volumetric soil is saturated. At the same time, there is a direct feedback between evaporation and soil moisture with no hydraulic diffusion process controlling the water movement inside soil. In this regard, the bucket-type soil hydrological model tends to simulate a rapid response of bare soil wetness to changes in atmospheric forcing. Meanwhile, Scott et al. (1997) found that in regions with dense canopy coverage, bucket models showed relatively slow response of evapotranspiration to precipitation forcing as there is a lack of canopy interception component in such models to reflect the rapid canopy transpiration of intercepted water.

An intermediate scheme in soil moisture simulation is the so-called force-restore model in which there is a thin top layer and a deep soil layer. Soil moisture in these two layers is governed by

$$\frac{dw_1}{dt} = -C_1 \frac{E_g - P}{rd_1} - C_2 \frac{w_1 - w_{ge}}{t} \tag{2}$$

$$\frac{dw_2}{dt} = -\frac{E_g + E_{tr} - P}{rd_2} \qquad (3)$$

where w_1 and w_2 are the soil moisture in the upper and deep layers. d_1 and d_2 are the depths of these two soil layers, E_g and E_{tr} represent ground evaporation and canopy transpiration, w_{ge} is the soil moisture when gravity force balances soil capillarity force, and J is the force-restore time scale. In this scheme, two different time scales are involved in representing the feedbacks between land and the atmosphere: there is a rapid response to the atmospheric forcing in the top thin layer and a slow restore process in the deep layer by soil moisture supply from deep soil to the upper layer for surface evaporation. Surface runoff occurs when the upper layer is saturated even though the deep layer may still be unsaturated and a contribution from canopy transpiration is included in the total evapotranspiration.

The opposite extreme from the "bucket" is a multi-layer soil model. This type of scheme has been developed to fully couple soil hydraulic diffusion processes with canopy and root-zone processes of water flow in the soil.

$$\frac{\partial q}{\partial t} = -\frac{\partial q}{\partial z} + S \qquad (4)$$

where 2 is soil water content, q is Darcian soil water flux density and S is a sink term which includes root water uptake affected by root density and distribution, canopy transpiration and in some schemes even horizontal runoff.

Comparing the three approaches to simulating soil moisture in GCMs, it is clear that different complexities in the land-surface parameterisation may affect model predictability. For instance, considering how soil moisture responses to meteorological forcing with bare soil surface condition:

- In single-layer bucket-type models: volumetric soil moisture depletes (fills) immediately when evaporation (precipitation) occurs. This leads to a rapid response from the land-surface to any atmospheric forcing and a short "memory" of soil moisture in the climate system and, ultimately, a less predictable land-surface component in the overall model.

- In force-restore and multi-layer scheme: deep soil water depletes (fills) by first diffusing through upper-layer and then evaporation from (after precipitation onto) surface. This leads to slower restoring responses within the soil layers, and thus to a longer "memory" of land-surface processes in the climate system, ultimately to a more predictable component in the overall model.

In recent years, there have been a number of studies focused on assessing the degree to which land-surface parameterisations can affect the predictability of climate models. Koster and Suarez (1996) investigated the influence of soil moisture retention on precipitation statistics. Scott et al. (1997) studied the timescales of land-surface evapotranspiration responses in the land-air feedback processes. Recently, Koster and Suarez (2001) reported on a study of soil moisture memory in climate models by constructing a soil moisture auto-correlation equation with components representing the nonstationary effects of atmospheric forcing, evaporation, runoff, and the correlation of atmospheric forcing with soil moisture condition. As AMIP2 models use a wide range of land-surface parameterisations, analysing results from AMIP2 simulations may offer some insight to understanding the relationships among physical processes parameterised.

This report presents preliminary analysis of sixteen AMIP2 model experiments. As a first step towards understanding the impacts of land-surface modelling on GCM simulations, the current report serves primarily to demonstrate differences among model simulations. The explanation of these differences is currently under intensive research and will be reported later. This report is structured as follows: Section 2 describes the AMIP2 models in this study and particularly the land-surface components in these models. In addition, validation datasets used in this analysis are introduced here. The skill of the models in simulating surface climatologies is described in Section 3. Section 4 presents the model skill in simulating the variations of surface climate and fluxes. Section 5 explores connections between model skill in simulating climate and surface flux anomalies. The results of lag correlation analyses are presented in Section 6 to initiate examination of whether land-surface parameterisations in GCMs affect model predictability. Finally, discussions and preliminary conclusions from the current analysis are presented in Section 7.

2. Models and Validation Data Used in the Analysis

Seventeen-year (1979-1995) AMIP2 simulations from sixteen AGCMs have been released at the time of this report. Table 1 lists the models used in the study, which include well-known research institutes and organisations. However, the models will remain anonymous in the rest of the analysis. Table 2 summarises some key aspects of the model configurations which are directly related to the focus of this study. It should be emphasised here that the purpose of this study is not to identify the best model(s), rather it is dedicated to improving understanding of model performance following the model physics. As shown in Table 2, there is a great variety of land-surface complexities in these models, ranging from simple Manabe-type bucket models with no

canopy related processes, intermediate bucket plus simple canopy stomatal resistance schemes, and finally, schemes with fully parameterised canopy processes and incorporated carbon-cycles. The soil hydrology component varies from the simple one-layer bucket scheme, through two-layer force-restore approaches to multi-layer soil schemes. Such diversity in complexity in land-surface schemes provides a good opportunity to study the role of land-surface parameterisation in climate simulations.

As noted by Phillips et al. (2000), the lack of high quality and globally consistent continental-scale observations of land-surface variables, such as surface heat fluxes, soil moisture and runoff, hampers the evaluation of GCM model performance. This study uses available observational datasets and model-derived reanalysis datasets while recognising the deficiencies in both types of validation data (e.g., Irannejad, et al., 2001). The Australian Bureau of Meteorology (BoM) observed rainfall and temperature datasets are employed for validation of model simulated surface climate. These latter data, originally formed on 0.25^o by 0.25^o grids, have been transformed to the common T62 grids to which all the sixteen AMIP2 models have been regridded. The BoM evapotranspiration climatology (Wang et al., 2001) is used in the evaluation of surface evaporation simulations. These data have been derived using the formulation of Morton (1983) and observed meteorological forcing (precipitation, temperature, radiation etc.), together with observed river discharge and observed precipitation in the calibration designed to conserve surface water balance. The NCEP/DOE AMIP2 reanalysis (http://wesley.wwb.noaa.gov/reanalysis2/) for the period of 1979 to 1995 is also used in later part of the study which examines the correlation between model skill in simulating surface climate anomalies and skill in simulating surface flux anomalies.

Besides the calculation of standard measurements, such as root-mean-square-error (RMSE), biases and spatial and temporal correlations between model results and observations, Linear Error in Probability Space (LEPS) skill score (Potts et al., 1996) is also used in this study. This score is related to the difference between the position of the simulation and the observation in the cumulative probability distribution space of the particular climate variable under consideration. This skill score has been used in the verification of the BoM statistical seasonal forecasting system (e.g., Jones, 1998; Drosdowsky and Chambers, 2001) and in the assessment of the BMRC experimental AGCM seasonal forecasts (Frederiksen et al., 2001).

For individual simulations, if the position of the simulation in the cumulative distribution in the model is P_f (ranging from 0.0 – 1.0) and that of the observation is P_v, then the LEPS skill score (Potts et al., 1996) is defined to be

$$S = 3(1 - |P_f - P_v| + P_f{}^2 - P_f + P_v{}^2 - P_v) - 1. \tag{5}$$

To achieve a skill score range from 100% to –100%, average skill (SK) may be defined (Potts et al., 1996) as

$$SK = \frac{\sum 100S}{\sum S_m}. \tag{6}$$

Here the summation is over all pairs of simulations and observations, where S is the individual score for each forecast, and S_m depends on whether the numerator is positive or negative. For a positive numerator, S_m is the sum of the maximum possible scores given the observations (obtained by setting $P_f = P_v$ in Equation 5. If the numerator is negative, S_m is the sum of the moduli of the worst possible scores given the observations, obtained by setting either $P_f = 1.0$ or

$P_f = 0.0$ in Equation 5 and taking the negative value with the largest modulus. As shown by Potts et al. (1996), the relationship between LEPS score and correlation coefficient is nonlinear, which is a function of standard deviation of model simulations: a LEPS score of zero corresponding to zero correlation and a LEPS score of 30% roughly corresponding to a correlation coefficient of 0.5.

Considering that there likely are to be large differences between the cumulative probability distribution derived from the model simulations and the observations, we calculate P_f and P_v using the model-simulated cumulative probability distribution (derived from the model 17-year AMIP2 integration) and observed cumulative probability distribution (derived from 50-year observations from 1950 to 1999), respectively. This is different from Frederiksen et al (2001) in which both model simulations and observations are referred to observed cumulative probability distributions. As discussed in Potts et al. (1996), the approach used in this study is more appropriate for estimating the potential predictability in the models.

3. Surface climatologies of sixteen AMIP2 models

At first, we briefly discuss some fundamental features of observed surface climatologies over the Australian region. These climatologies will then be used in evaluating model simulated climatologies. Figure 1 shows the surface climatologies using the BoM's observational datasets. Rainfall, monthly mean daily maximum surface temperature (Tmax) and monthly mean daily minimum surface temperature (Tmin) climatologies are the averages over the period of 1950 to 1999, while surface evaporation climatology is from Wang et al. (2001). Figures 1(a) and (b) display significant seasonal variations of rainfall distributions in DJF and JJA. In the austral

summer, the Australian monsoon produces heavy rainfall over the northern and eastern part of the continent. Precipitation is low in the central, western and southern regions. In JJA, large rainfall is observed in the southern and western parts of the region, and precipitation is low in the central and north. Both Tmax and Tmin exhibit significant shifts seasonally following the sun, and it is generally cooler in the south and warmer in the north and central regions. Surface evaporation climatologies in DJF and JJA demonstrate a dry feature of the continent, with low surface evapotranspiration over a large part of the continent in both seasons. Surface evaporation exceeding 80 Wm^{-2} only occurs over the northern and eastern parts of the continent in the summer monsoon rainy season.

Figure 2 shows the biases of seasonally averaged precipitation climatologies from the sixteen AMIP2 models against the observed climatology for the period 1950 to 1999 from the BoM data. In the austral summer season (DJF), about half of the models overestimate the rainfall climatology over a large part of the continent, with models B, F, N and P having greater than 50 mm $month^{-1}$ positive biases over a large part of the continent. Models A and D have systematic negative biases over the whole continent. In JJA, precipitation biases are much weaker in most of the models compared with DJF. Except for models E, M and O, most of the models show positive biases in the eastern part of the continent and most models exhibit negative biases in the south.

The area-averaged root-mean-square-error (RMSE) of the model climatologies over the continent is shown in Figure 3. Simulations from a "poor-mans ensemble", generated by simply averaging all the model simulations throughout the 17-year period as a set of new simulations, are also included for comparison. In agreement with results from Figure 2, RMSE is generally higher in the austral

summer season, with model P being the significant outlier. As found by Ebert (2001) and Love et al. (1995), such a simple averaging all the model simulations (poor-man's ensembles) gives the best overall RMSE results.

Despite the large biases seen in Figure 2, the spatial pattern of rainfall climatology and its seasonal migration associated with the Australian monsoon system are reasonably well simulated by most of the AMIP2 models (Figure 4). The spatial correlations with observed rainfall climatology are around 0.8 for most models in the summer season, but the correlation drops to about 0.6 in the austral winter season. Of the sixteen models, C and P have the largest seasonal variations of the spatial correlation, with models C and P giving the poorest simulation of winter rainfall distribution.

Climatological biases of monthly averaged daily maximum surface temperature (Tmax) between model simulations and the BoM observations are shown in Figure 5. In the austral summer season (Figure 5a), the bias can be as large as 5 to 10 ^{o}C. Most models underestimate the daily Tmax in the northern and eastern parts of Australia with models F, G, N, O and P having negative biases of above 5 ^{o}C. Most models also tend to overestimate Tmax in the southern part of the continent by 2 to 5 ^{o}C. In contrast, models A, H and M show systematic positive biases over the whole continent and models G and O have systematic negative biases. In the austral winter (Figure 5b), most models tend to underestimate Tmax except models A, B, E, H and M, which have positive biases over much of the continent. Models F, G, N and O have systematic negative biases over almost the whole continent, with the largest biases seen in Model O of 5 to 10 ^{o}C. Referring to Table 2, results

suggest that models with a lower number of levels in soil temperature calculations than the average are likely to be part of the outliers.

Area-averaged RMSE in Figure 6 demonstrate that the models have relatively low errors in the winter Tmax simulations and, again, the poor-man ensembles outperform most models. Figure 6 shows that models H and O have larger RMSE than the others, with a large seasonal feature in model H. In contrast to the results seen in RMSE, results of spatial correlations with observed Tmax climatology (Figure 7) show that the pattern of Tmax distribution in most models are similar to the observed, particularly in the winter season. There are large model discrepancies in the results from October to March, with models H, N, O and P showing lower skill. Together the results in Figures 5 to 7, tends to suggest that models O and P (with one-layer bucket type land-surface schemes) exhibit poorer performance in the simulation of Tmax climatology than the others. The poor performance of model H is also consistent with the relatively low number of soil layers in its calculations of surface energy balance (ref. Table 2).

Figure 8 shows the climatological biases between model simulated and observed daily minimum surface temperature (Tmin) in DJF and JJA. In DJF, some models (e.g. A, D, F, J and M) show similar features to their Tmax simulations while others exhibit many differences. For instance, models G, H and K show almost opposite biases from those in Tmax. Model H has positive biases in Tmax and negative biases in Tmin, resulting in significant overestimation of daily temperature variation in DJF (Figure 11a). In contrast, models G and O have lower Tmax and higher Tmin, leading to the underestimation of daily temperature variations in this season. Except for some outliers, results from daily temperature range (Tmax-Tmin) in Figure 11a suggest that the biases of

surface temperature daily range are between 2 to 5 ^{o}C, smaller than the biases in Tmax and Tmin individually.

Similar results are seen in Tmin simulations in JJA with models G, H and O showing biases opposite to those in Tmax. Consequently, the daily temperature range is poorly simulated in these models (Figure 11b). Area-averaged RMSE (Figure 9) shows that models B, G, H and N have larger errors than most of the others in simulating Tmin climatology. However, models C, L, N, O and P generate relatively poorer simulations of the spatial distributions of Tmin in Australian region (Figure 10). Combining results from Figures 5 and 11, it seems that models O and P (with one-layer formula for the surface energy balance), together with models having lower number of layers in calculating surface temperature (e.g., models G and H), are more likely poorer at simulating the Tmax and Tmin climatological distributions in the season from October to March. In addition, models C and L exhibit different capability in Tmax and Tmin simulations.

Using the BoM evapotranspiration climatology data from Wang et al. (2001), the sixteen AMIP2 model-simulated surface evaporation climatologies are evaluated in Figures 12 to 14. In DJF, there are quite large differences among the model simulations, with most models showing 20 to 50 Wm^{-2} biases over the continent. Models A, D, I and M have large negative biases over much of the continent, while, B, F, N and P show systematic positive biases in DJF. In addition, the majority of the models underestimate surface evaporation over the central-western dry region. Comparing these conclusions with the Tmax results in Figure 5a, tends to suggest that when surface evaporation is underestimated, the models have positive biases in Tmax simulations and vice versa. However, some models do exhibit different features. For instance, model B overestimates evaporation over

much of the continent but it shows negative biases over the north. Similarly, model H has positive biases over the whole continent in Tmax simulations but underestimates surface evaporation in the north. Similarly, the relationship between evaporation and Tmin simulations varies with models. There is no obvious coherence between the biases in surface evaporation simulation and Tmax and Tmin simulations, and some models show closer linkage than others.

In JJA, the biases simulated in most models tend to be weaker than in DJF, a feature also seen in the area-averaged RMSE in Figure 13. For instance, models B, F and N show much weaker positive biases in the simulations of winter surface evaporation climatology. Large biases in the coastal region in models K and L occur in both JJA and DJF. Such biases contribute to the large RMSE results in Figure 13. In addition, model P exhibits large RMSEs in most of the months except in the winter season. The pattern of evaporation climatology is reasonably reproduced by a large number of models (Figure 14), with higher correlation in the summer than in winter season. However, there are also a number of models, in particular, models K, L, M and P, which have low skill in simulating the spatial distribution of surface evaporation.

Overall, in this section, the performance of sixteen AMIP2 models in simulating surface climatology over the Australian region has been evaluated by the calculation of climatological biases, RMSE and spatial correlations. In the following section, the model skill in simulating climate variations over this region is assessed.

4. Results of the AMIP2 models in simulating surface climate anomalies

One of the purposes of this study is to assess the predictability of climate anomalies over this region. To this end, the skill of each model in simulating monthly precipitation and Tmax and Tmin anomalies for the 17-year AMIP period is presented here.

Figure 15 shows the model skill in simulating monthly precipitation anomalies against the BoM observational data. Model simulated anomalies are relative to the 17-year climatology from the model simulations. Also, the cumulative probability distribution of such anomalies derived from each model is used in the LEPS calculations, together with the distribution derived from the observations. As shown in Figure 15a, most of the models exhibit rather limited skill in simulating rainfall anomalies in DJF. There is also no coherence about regions where most models have (or do not have) skill. The majority of the models do not show skill in simulating rainfall variations in the north and northeast where the Australian summer monsoon dominates. Part of the skill seen in the central region may be due to the consistently dry climate here. In the austral winter season (JJA), the model skill in simulating rainfall anomalies is better than that in DJF. This is particularly clear in models B, D, F and M. Area-averaged LEPS score (not shown) also suggest more skill in the winter season, as do statistical seasonal forecasting systems (e.g., Drosdowsky and Chambers, 2001). In this season, models D, F and M outperform most others in simulating winter rainfall variations.

Even though there is little or no coherence in the model LEPS score of Tmax simulations, the spatial distribution of each model's LEPS score is much smoother (Figure 16) than the model skill for rainfall variations. The BoM statistical seasonal forecasting system (e.g., Jones, 1998) has reasonable skill in Tmax forecasting over the eastern part of the continent in DJF, but this is not commonly seen in the AMIP2 model simulations (Figure 16a), of which only half of the models (A, E, G, J, L, M and N) show skill in this region. The model skill is moderately improved in JJA (Figure 16b) with models A, D, F, and M exhibiting reasonable skill over much of the continent. This is in agreement with the BoM statistical model, which also shows skill over extensive areas of the continent in JJA (Jones, 1998).

In the summer season (DJF), results from Figure 17a show that the skill of most of the models in simulating Tmin variations has broadly similar features to those seen in Tmax. These similarities are not retained in the winter season (Figure 17b). For instance, models A and F have skill in Tmax simulations in the eastern part of the continent, while rather poor model skill is seen in the Tmin simulations for the same region. Such differences are, to some extent, also found in the skill of statistical models in forecasting Tmax and Tmin anomalies over the Australian region (Jones, 1998).

As there are no multiyear surface evaporation data available, only the climatology used in Section 3, we cannot calculate the LEPS score of each model in simulating surface evaporation variation over the Australian region. However, in the next section, such skill will be assessed by comparison with the NCAR/DOE reanalysis datasets.

5. Linkage between model skill in simulating anomalies of surface climate and fluxes

In Section 4, we have assessed of the performance of the AMIP2 models over the Australian region in terms of the model skill in simulating surface climate variations from the model 17-yr integrations. We next examine whether the skill of the models in simulating surface climate is linked to the skill of the models in simulating surface fluxes. If meteorological forcing, such as radiation, rainfall, temperature etc, is the dominant factor in the model simulation of surface fluxes, then good correlations between the LEPS score of surface climate and LEPS of surface fluxes in the model simulations would be expected. In contrast, low correlations would suggest that differences in complexity and approach used in representing the land-surface processes are important for air-land interaction simulations.

As there are no multi-year observed surface flux datasets available, the 17-yr NCAR/DOE reanalysis is employed in calculating the LEPS score of each model's surface latent and sensible heat fluxes. To be consistent, each model's LEPS score in rainfall and surface temperature simulations is re-calculated by using the same reanalysis data. Correlations of the LEPS scores of the models in simulating surface climate anomalies (temperature, precipitation and soil moisture) and surface flux anomalies (latent and sensible heat fluxes) are also analysed. It should be noted that there are deficiencies in the reanalysis data (e.g., Irannejad et al., 2001) and that there are incompatibilities of some fields (e.g., total soil moisture) between models and reanalysis. However, as emphasis is on the connections between the LEPS scores rather than on the LEPS scores themselves, use of the reanalysis data for this part of calculation is reasonable (even though not desirable). In addition, applying both the cumulative probability distributions derived from model

simulations and from the reanalysis data in the LEPS calculations further alleviates the impacts of the deficiencies in the reanalysis dataset on this part of the study.

Figure 18a shows the correlation of LEPS scores in the model simulations of precipitation anomalies and surface evaporation anomalies over the 17-yr AMIP2 period for the 16 AGCMs. The averaged correlation coefficient (heavy dashed line) is about 0.3 to 0.6, with large variations between models. Figure 18a suggests that the connection between these two skills in model simulation is more pronounced in the austral summer season. This feature is clearer in the results from the poor-man ensembles (heavy solid line). To illustrate this feature further, Figure 19 presents scatter plots of rainfall and surface evaporation LEPS scores over the Australian region from one of the models. In the summer season, correlation between the two LEPS scores can be seen. When the model LEPS score for simulating the precipitation anomalies is high, there also is a high LEPS score in simulating surface evaporation anomalies.

As AMIP2 results are from experiments that include complex coupled land-atmosphere interactions, it is difficult to identify cause and effect relationships. For example, does a good simulation of precipitation lead to more skilful simulation of surface evaporation, or, does successful representation of surface evaporation (capturing surface water recycling) contribute to the model skill in rainfall simulations? Detailed, process analysis (to be reported separately) suggests that during the summer season, surface radiative forcing is sufficient in this region and as a consequence, surface water availability constrains the surface evaporation processes in most models. Therefore, it is reasonable to attribute higher LEPS score in a given model surface evaporation simulation to skilful simulation of precipitation anomalies. In contrast, during the

20

austral winter (Figure 19b), there is virtually no correlation between the skill of the model in simulating rainfall anomalies and skill in simulating surface evaporation anomalies. Thus, even if a model has reasonable skill in simulating rainfall anomalies, this does not deliver skill in simulating surface evaporation anomalies. These results suggest that in the winter, surface flux simulations are not dominated by the meteorological forcing simulated in the model. Model representation of surface energy and water partitions plays a significant role in simulating the variation of surface fluxes in this season. It should be pointed out that such relationships seen over the Australian continent might be different from results in other regions with different climates. This will be studied by further analysis of model results over different regions in the future.

In contrast to the relationships seen in Figure 18a, the model skill in simulating soil moisture anomalies derived from the reanalysis data is only loosely correlated with precipitation anomalies from the reanalysis (Figure 18b). As discussed before, soil moisture in the reanalysis may be quite different from the models. Nevertheless, our interest here is not on the model skill of reproducing the soil moisture variations in the reanalysis, but on the connection between LEPS scores in rainfall and soil moisture simulations. The scatter among the models in Figure 18b is also larger than seen in Figure 18a. These results suggest that variations of soil moisture in the models are not solely determined by rainfall simulations. Rather, soil moisture seems to be determined by surface hydrological process parameterisation in the models. In addition, results here indicate that improvement and widened application of numerical model forecasts of climate anomalies will require better treatment of hydrological processes in models. For instance, soil moisture conditions are more important for agriculture than rainfall anomalies and, it seems, forecasts of rainfall anomalies do not necessarily provide any indication of soil wetness.

6. Potential impacts of land-surface modelling on model predictability

As discussed in the introduction, this study is, in part, dedicated to investigating the potential impacts of land-surface modelling on GCM predictability. A number of studies have reported recently on this topic (e.g., Koster and Suarez 1996; Scott et al. 1997; and Koster and Suarez 2001). Here, a preliminary analysis of the impacts of land-surface modelling on model predictability by lag-correlation calculations is presented. The seasonal cycle is removed before the calculation of lag-correlation between two variables. With 204 samples from the 17-yr model integrations used in the calculation, a correlation coefficient exceeding about 0.14 is statistically significant with 95% confidence (assuming the samples are independent). Lag correlations up to twelve months are calculated in this study.

Figure 20 shows the three-month lag-correlation between soil moisture (mrso) and surface evaporation (hfls) over the Australian region, with surface evaporation lagging soil moisture. A three-month lag is chosen here as this is the timescale reported by most AGCM seasonal forecasts (Frederiksen et al., 2001). There are remarkable differences among AMIP2 models in terms of the timescale over which soil moisture anomalies affect surface energy partitions. Lag correlations between these two variables are, overall, positive over the continent, except for model C which has negative correlations over part of the region. Among the 16 models, it is noted that models O and P, together with models C, M and N, exhibit the lowest overall lag correlations. As will be shown later (ref. Figure 25), such results from models O and P are directly due to the short retention period of soil moisture anomalies in the land-surface schemes used in the models with a bucket-type structure.

The three-month lag correlations between soil moisture and precipitation anomalies (not shown) are substantially lower than the results seen in Figure 20. There is no significant correlation in any of the 16 AMIP2 models considering a three-month lag. This is chiefly the result of other processes affecting the simulation of precipitation in these AGCMs. However, soil moisture anomalies do affect surface temperature forecasts. Figure 21 shows that with a three-month lag, anomalous soil moisture conditions are linked to the forecasts of surface temperature anomalies in a number of models, particularly over the eastern part of the continent. Among the 16 models, models A, G and K have the largest area of significant correlations and, again, models O and P with a bucket-type land-surface scheme, as well as C, L and N, show lowest lag correlations.

The characteristic feature, exhibited by models O and P, can be largely explained by the short retention time for soil moisture anomalies when simulated by simple bucket-type surface schemes. Figure 22 displays the auto-correlation of soil moisture anomalies across all the AMIP2 models. Clearly, models O and P have the lowest auto-correlation on this time scale. Combining results from Figures 20 to 22, it is possible to draw two complementary conclusions: (i) in models with a simple one-layer bucket land-surface scheme, the retention of soil moisture anomalies is much shorter than others, resulting in a weak influence from land-surface conditions on the model predictability on a seasonal time scale; (ii) predictability in GCM models is also affected by factors other than the contribution from land-surface processes. For instance, model C shows high auto-correlation of soil moisture anomalies in Figure 22, but it has consistently lower lag correlations in Figures 20 and 21, implying other factors in the model are contributing to this model's low

predictability. Jointly, these results prompt the need for further research in an effort to more fully understand the mechanisms involved.

To further illustrate the differences among simulations of soil moisture variations simulated in the AMIP2 models, Figure 23 shows the soil moisture anomalies from three models (D, J and O) after the removal of their simulated seasonal cycles. Results from the grid point located near 25^oS and 135^oE are analysed here. These three models have been selected to encompass the range of simulated characteristics of interest in this diagnostic subproject. Model D represents models having high soil moisture auto-correlations, model J is typical of models with modest auto-correlations, and model O illustrates results from models with much lower auto-correlations, as seen in Figure 22.

These distinguishing features are clearly shown in Figure 23. Model D exhibits a slow variation of soil moisture anomalies and it also shows a slow downward trend simulated in soil water which contributes to the large auto-correlations. Model J shows moderate variations of soil moisture conditions at this location, while model O with a bucket-type scheme exhibits rapid responses to the meteorological forcing, with soil moisture anomalies responding rapidly to rainfall anomalies and evaporative demand (not shown). Thus, for this model, an anomalous wet condition decays rapidly over a short time scale.

To illustrate the time scale of soil moisture processes simulated by each of the models, Figure 24 shows areally-averaged lag-correlations from zero to twelve months lag. Results are area-averages over the Australian continent. Figure 24a shows large differences in models' lag correlations

between soil moisture and evaporation. Some AMIP2 models show averaged correlations exceeding 0.15 up to five and six months lag, while others exhibit low correlations within a month. These differences are even more pronounced in the correlations between soil moisture and surface temperature (Figure 24b).

The rapid decaying of correlations in models O and P seen in Figure 24 is consistent with the features of soil moisture anomalies simulated and already described for these two models (Figure 25). Figure 25 demonstrates the remarkably different characteristics of soil moisture variations simulated in these two models as compared with the other AMIP2 models. These two models have the fastest decaying rates of soil moisture auto-correlations, underlining the short retention time of soil moisture anomalies and subsequently other surface climate anomalies. For example, Figure 25b shows that the retention of surface temperature anomalies in models O and P is shorter than for most other models examined here.

Overall, the lag and auto-correlation analysis of AMIP2 models clearly shows the potential of land-surface schemes to influence model-simulated characteristics of surface climate anomalies. The impacts of different complexity of land-surface modelling on the model predictability have also been illustrated, particularly by comparing the behaviours of AGCMs coupled to the simplest (bucket-type) land-surface scheme with the performance of models incorporating more complex surface representations.

7. Discussion and Conclusions

This report has presented the results of analysing 16 AMIP2 models over the Australian region. The study has focused on assessing the surface climate and surface fluxes simulations in the current AGCMs forced by observed SSTs and sea-ice. The goals have been (i) to explore the impacts of land-surface modelling on the skill of surface climate simulations and (ii) to evaluate the potential of land-surface complexity to affect model predictability.

In this preliminary analysis, Australian Bureau of Meteorology (BoM) observational rainfall, temperature and surface evapotranspiration datasets have been employed to evaluate the surface climatologies simulated by the 16 models. The Linear Error in Probability Space (LEPS) score has been used to assess the skill of the models in simulating surface climate anomalies for the 17-year AMIP2 period (1979 to 1995). This study has quantified the skill of the model simulations using the measurements including biases, rmse and spatial correlations. A range of model differences have been described and linked to the complexity in the model's land-surface schemes. For instance, results tend to suggest that models using a small number of soil layers (1 or 2) in the calculation of surface energy balance generate poorer simulations of Tmax and Tmin over the region than those with more soil layers. The connection between model skill in simulating surface climate anomalies and surface flux anomalies has been explored. It is found that in the austral summer, the skill of models in simulating precipitation anomalies is correlated to their skill in simulating flux anomalies. However, these correlations become weaker in the winter season, implying that the influence of land-surface modelling is more significant then.

Lag-correlation analysis has revealed that the characteristics of climatic "memory" from land-surface processes (e.g. soil moisture) differ among the sixteen models: some models show rapid feedback processes between the land-surface and the overlying atmosphere, while others exhibit slowly varying processes in which anomalous surface conditions have influences over longer time periods. It was found that models with simple bucket-type schemes tend to show rapid decay rates in soil moisture anomalies, leading to much weaker lag correlations between soil moisture conditions and surface climate anomalies. Overall, this analysis suggests that land-surface modelling has the potential to affect AGCM predictability on seasonal and even longer time scales. This impact on the predictability skill is itself a function of the land-surface scheme characteristics.

It should be noted that the current study is only a first step towards a better understanding of the role of land-surface modelling in climate simulations. This report has quantified model skill and explored differences among models and discrepancies between simulations and observations. Ongoing research will continue to try to determine why the models are different, and to what extent different model performances can be linked to the complexity of their land-surface representations. More detailed analyses such as those reported in Koster and Milly (1997), Gedney et al. (2000) and Koster and Suarez (2001) will be pursued in the future. With the continuing release of AMIP2 model results, including those from the BMRC Atmospheric Model (BAM) AMIP2 experiments, analysis will become more inclusive and more geographically extensive.

Acknowledgment:

References:

Drosdowsky, W., and L. Chambers, 2001: Near-global sea surface temperature anomalies as predictors of Australian seasonal rainfall, *J. Climate*, **14**, 1677-1687.

Ebert, E.E., 2001: Ability of a poor man's ensemble to predict the probability and distribution of precipitation. *Mon. Wea. Rev.* (in press).

Frederiksen, C.S., H. Zhang, R.C. Balgovind, N. Nicholls, W. Drosdowsky and L. Chambers, 2001: Dynamical seasonal forecasts during the 1997/98 ENSO using persisted SST anomalies, *J. Climate*, **14**, 2675-2695

Gates, W.L., J.S. Boyle, C. Covey, C.G. Dease, C.M. Doutriaux, R.S. Drach, M. Fiorino, P.J. Gleckler, S. Marlais, T.J. Phillips, G.L. Potter, B.D. Santer, K.R. Sperber, K.E. Taylor, and D.N. Williams, 1999: An overview of the results of the Atmospheric Model Intercomparison Project (AMIP I). *Bull. Amer. Meteor. Soc.*, **8**0, 29-55.

Gedney, N., P.M. Cox, H. Douville, J. Polcher, and P.J. Valdes, 2000: Characterising GCM land surface schemes to understand their responses to climate change, *J. Climate*, **13**, 3066-3079.

Henderson-Sellers, A., K. McGuffie, and A.J. Pitman, 1996: The Project for the Intercomparison of Land-surface Parameterization Schemes: 1992-1995. *Climate Dyn.*, **1**2, 849-859.

Henderson‑Sellers, A., 1999: Atmospheric global climate models' representation of the land surface: From AMIP I to AMIP II.

Accessible at http://www-pcmdi.llnl.gov/pilps3/amip1review/amip1review.pdf

Irannejad, P., A. Henderson‑Sellers, Y. Shao, and P.K. Love, 1995: Comparison of AMIP and PILPS off‑line landsurface simulations. Proceedings of the First AMIP Scientific Conference, 15-19 May 1995, Monterey, California, WMO/TD‑No. 732, pp. 465-470.

Irannejad, P., A. Henderson‑Sellers, T.J. Phillips, K. McGuffie, A.J. Pitman, and E.J. Clayton, 2001: Evaluation of the continental component of the global water cycle: The dearth of validation data, the 4th International Scientific Conference on the Global Energy and Water Cycle, 10-14 September 2001, Paris, France.

Jones, D.A 1998: The prediction of Australian land surface temperatures using near global sea surface temperature patterns. *BMRC Research Report No. 70.* Bureau of Meteorology, Melbourne, Australia, 44pp.

Koster, R.D., and P.C.D. Milly, 1997: The interplay between transpiration and runoff formulations in land‑surface schemes used with atmospheric models, *J. Climate*, **10**, 1578-1591.

Koster, R.D., and M.J. Suarez, 1996: The influence of land surface moisture retention on precipitation statistics, *J. Climate*, **9**, 2551-2567.

Koster, R.D., and M.J. Suarez, 2001: Soil moisture memory in climate models, *J. Hydrometeo.*, **2**, 558-570.

Love, P.K. and Henderson-Sellers, A., 1994, Land surface climatologies of AMIP-PILPS models and identification of regions for investigation, PCMDI Report 12, Lawrence Livermore Laboratory, California, 83pp

Love, P.K., Henderson-Sellers, A. and Irannejad, P., 1995, AMIP diagnostic subproject 12 (PILPS Phase 3): land-surface processes, Procs. of The First International AMIP Scientific Conference (Monterey, CA, USA, 15-19 May 1995) (ed. W.L. Gates), WMO/TD-No. 732, WCRP, Geneva, 101-106.

Manabe, S., 1969, Climate and the ocean circulation I: the atmospheric circulation and the hydrology of the earth's surface. *Mon. Wea. Rev*., 97, 739-774.

Morton, F. I., 1983: Operational estimates of areal evapotranspiration and their significance to the science and practice of hydrology, *J. Hydro.*, **66**, 1- 76.

Phillips, T.J, and Co-authors, 2000: Validation of Land-Surface Processes in AMIP Models: A Pilot Study, PCMDI Report No. 63, PCMDI, Livermore, CA, USA, 22pp.

Phillips, T.J., A. Henderson-Sellers, P. Irannejad and H. Zhang, 2002: AMIP II diagnostic sub-project 12: Land-surface processes and parameterizations (a joint AMIP/PILPS project). Accessible at http://www-pcmdi.llnl.gov/amip/DIAGSUBS/SP12.html/.

Potts, J.M., C.K. Folland, I.T. Jolliffe, and D. Sexton, 1996: Revised "LEPS" scores for assessing climate model simulations and long-range forecasts, *J. Climate*, **9**, 34-53.

Qu, W.Q., and A. Henderson-Sellers, 1998: Comparing the scatter in PILPS off-line experiments with those in AMIP I coupled experiments. *Glob. Plan. Chang*e, **19**, 209-224.

Scott, R.L., D. Entekhabi, R. Koster, and M. Suarez, 1997: Timescales of land surface evapotranspiration reponse, *J. Climate*, **10**, 559-566.

Wang, Q.J., F.L.N. McConachy, F.H.S. Chiew, R. James, G.C. de Hoedt, and W.J. Wright, 2001, Climatic Atlas of Australia: Maps of Evapotranspiration. Available at http://www.bom.gov.au/climate/averages/climatology/evapotrans/text/et-txt.shtml.

Zhang, H., A. Henderson-Sellers, A.J. Pitman, J.L. McGregor, C.E. Desborough and J. Katzfey, 2001: Limited-area model sensitivity to the complexity of representation of the land-surface energy balance. *J. Climate*, **14**, 3965-3986.

Zhang, H., and C.S. Frederiksen, 2001: Impacts of soil moisture on AGCM seasonal forecasts: An experimental sensitivity study. *J. Climate* (under revision)

Zhang, H**.**, A. Henderson-Sellers, J.L. McGregor, and J.J. Katzfey, 2002: Complexity in land-surface parameterization and weather simulations: A case study of Tropical Cyclone Vance, *Mon. Wea. Rev.* (submitted)

Table Legend:

Table 1: Models used in the analysis

Model	Full name	Country
CCCMA	Canadian Centre for Climate Modelling and Analysis	Canada
CCSR	Center for Climate System Research	Japan
CNRM	Centre National de Recherches Meteorologiques	France
DNM	Department of Numerical Mathematics	Russia
ECMWF	European Centre for Medium-Range Weather Forecasts	UK
GLA	Goddard Laboratory for Atmospheres	USA
JMA	Japan Meteorological Agency	Japan
MRI	Meteorological Research Institute	Japan
NCAR	National Center for Atmospheric Research	USA
NCEP	National Centers for Environmental Prediction	USA
PNNL	Pacific Northwest National Laboratory	USA
SUNYA	State University of New York at Albany	USA
UGAMP	The UK Universities' Global Atmospheric Modelling Programme	UK
UIUC	University of Illinois at Urbana-Champaign	USA
UKMO	United Kingdom Meteorological Office	UK
YONU	Yonsei University	Korea

Table 2. Model codes and features of the sixteen AMIP2 models analysed in this report.

Code	Resolution	Land-surface components		No. of layers in soil temperature calculations	No. of layers in soil moisture calculations
		Soil model complexity	Canopy representation		
A	T42L18	bucket	const. canopy resistance	3	1
B	T63L45	force-restore	intercept. + transpiration	2	2
C	4x5 L21	multi-layer diffusion	intercept. + transpiration	24	24
D	T159L50	multi-layer diffusion	intercept. + transpiration	4	4
E	T63L30	multi-layer diffusion	intercept. + transpiration	4	3
F	T42L18	multi-layer diffusion	intercept.+ transpiration +CO_2	6	6
G	T62L18	multi-layer diffusion	intercept. + transpiration	3	2
H	T42L18	multi-layer diffusion	intercept. + transpiration	2	3
I	2.5x3.75 L58	multi-layer diffusion	intercept.+ transpiration +CO_2	4	4
J	2.5x3.75 L19	multi-layer diffusion	intercept.+ transpiration +CO_2	4	4
K	T47L32	multi-layer diffusion	intercept. + transpiration	3	3
L	4x5 L20	multi-layer diffusion	intercept. + transpiration	2	3
M	T42L30	multi-layer diffusion	intercept. + transpiration	3	3
N	T42L18	multi-layer diffusion	intercept.+ transpiration +CO_2	6	6
O	4x5 L24	bucket	no	1	1
P	4x5 L15	bucket	no	1	1

Figure Legend

Figure 1: Observed surface climatology over the Australian region from the BoM's observational datasets. (a) Precipitation climatology (mm month^{-1}) in DJF for the period of 1950-1999; (b) as (a) but for JJA; (c) Monthly mean daily maximum surface temperature climatology (^{o}C) in DJF for the period of 1950-1999; (d) as (c) but for JJA; (e) Monthly mean daily minimum surface temperature climatology (^{o}C) in DJF for the period of 1950-1999; (f) as (e) but for JJA; (g) Surface evapotranspiration climatology (W m^{2}) in DJF from the datasets of Wang et al. (2001); (h) as (g) but for JJA.
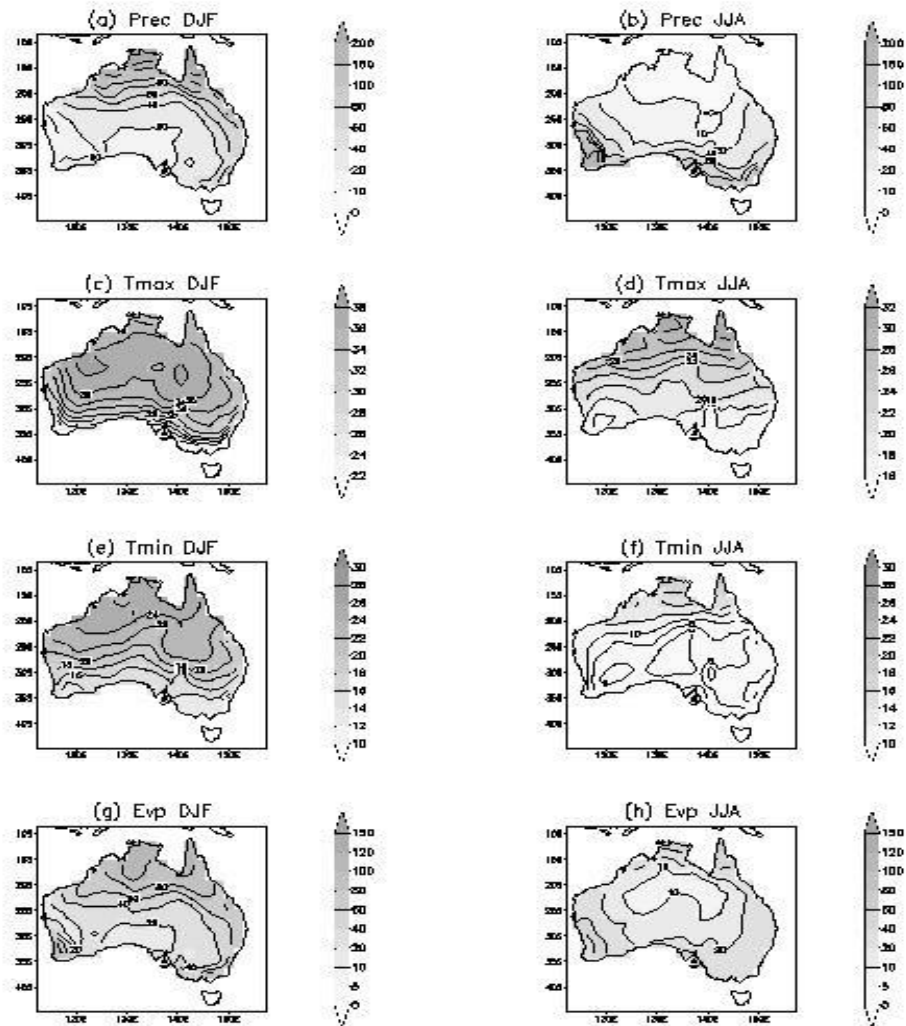
Figure 1: Observed surface climatology over the Australian region.

Figure 2: Climatological biases of precipitation simulated by 16 AMIP2 models against the BoM observations (mm month^{-1}).
(a) DJF;
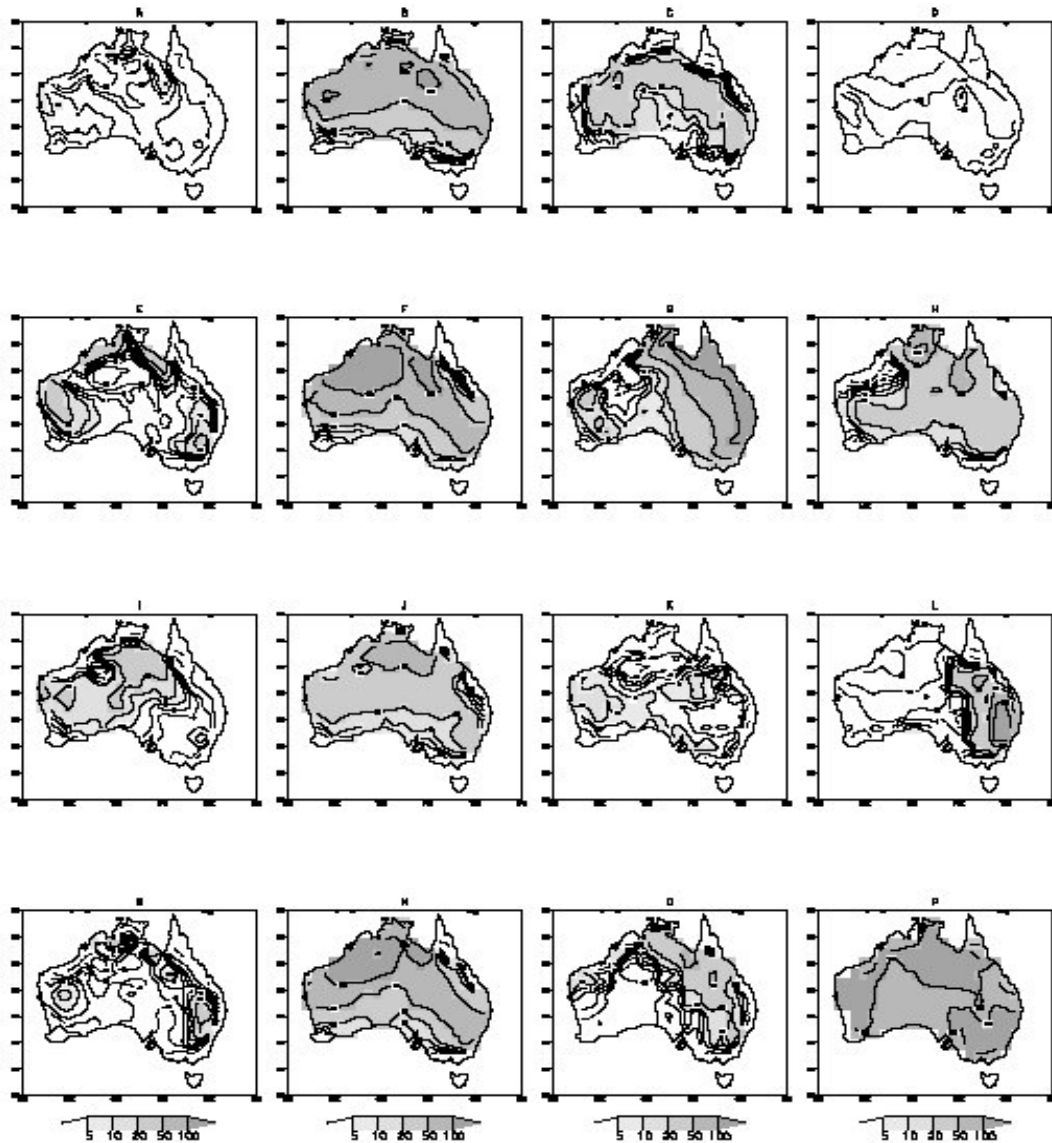
Figure 2(a): Precipitation bias in DJF. Unit: mm/month.

2(b) JJA.
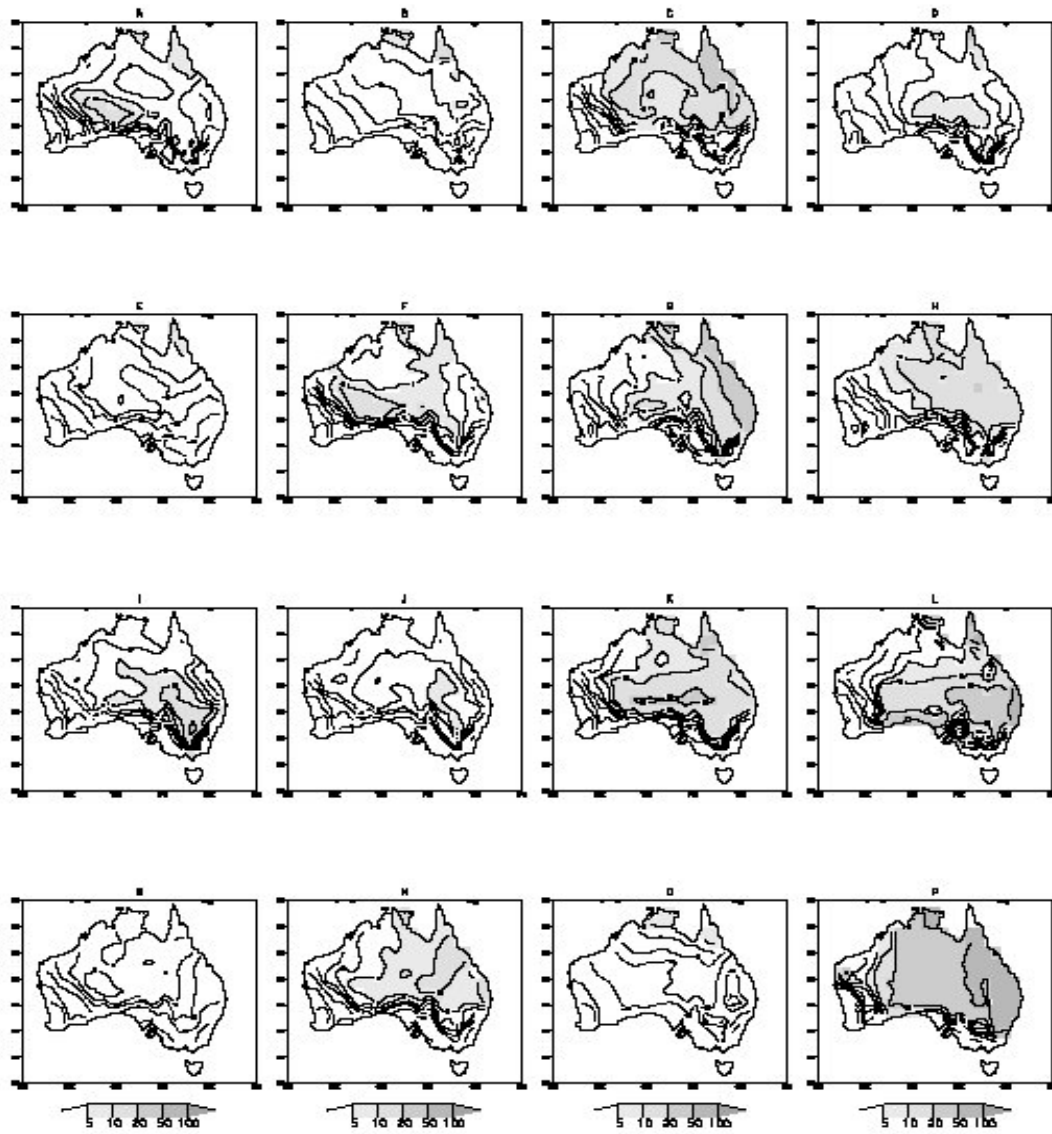
Figure 2(b): Precipitation bias in JJA. Unit: mm/month.

Figure 3: Root-mean-square-error (rmse) of precipitation climatology simulated by 16 AMIP2 models against the BoM observations (mm month[-1]) over the Australian region. The heavy dashed line represents the averaged rmse of all the 16 models. The heavy solid line represents the rmse of results from poor-man ensembles of the 16 models.
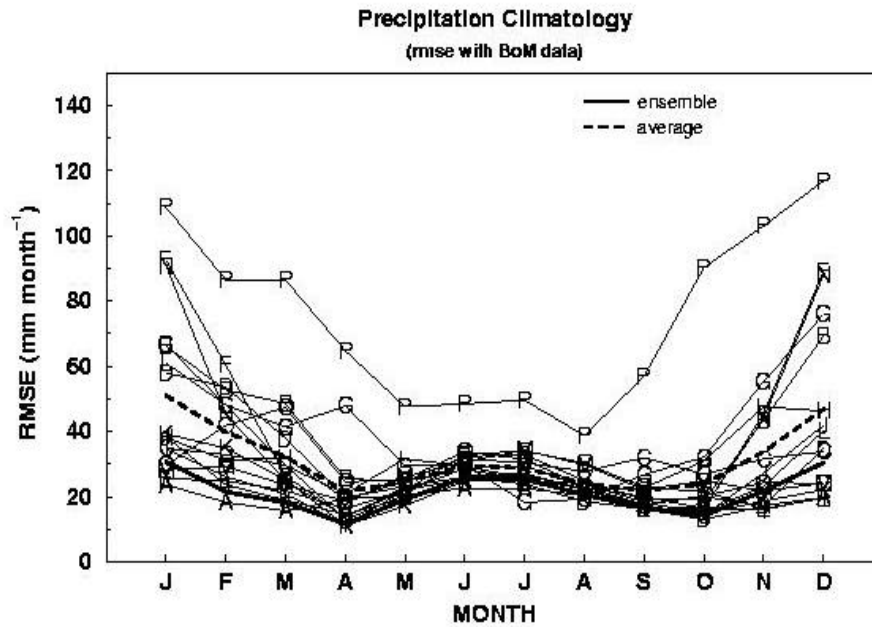
Figure 3: Root–mean–square–error (rmse) of precipitation climatology against the BoM observations (mm/month).

Figure 4: Spatial correlations of precipitation climatology simulated by 16 AMIP2 models against the BoM observations (mm month[-1]) over the Australian region. The heavy dashed line represents the averages of all the 16 models. The heavy solid line represents the results from poor-man ensembles of the 16 models.
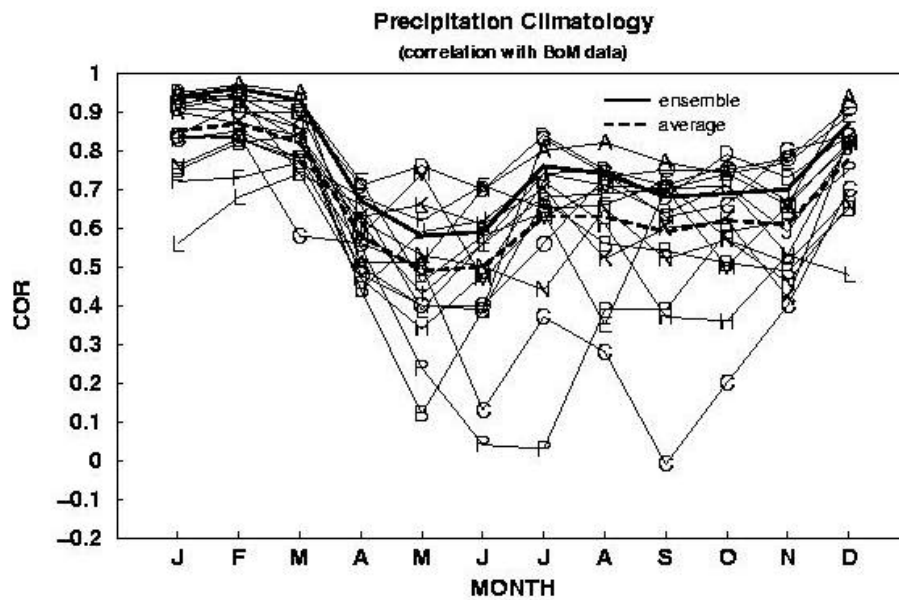
Figure 4: Spatial correlations of precipitation climatology against the BoM observations (mm/month).

Figure 5: As Figure 2 but for monthly mean daily maximum surface temperature (K).

(a) DJF;
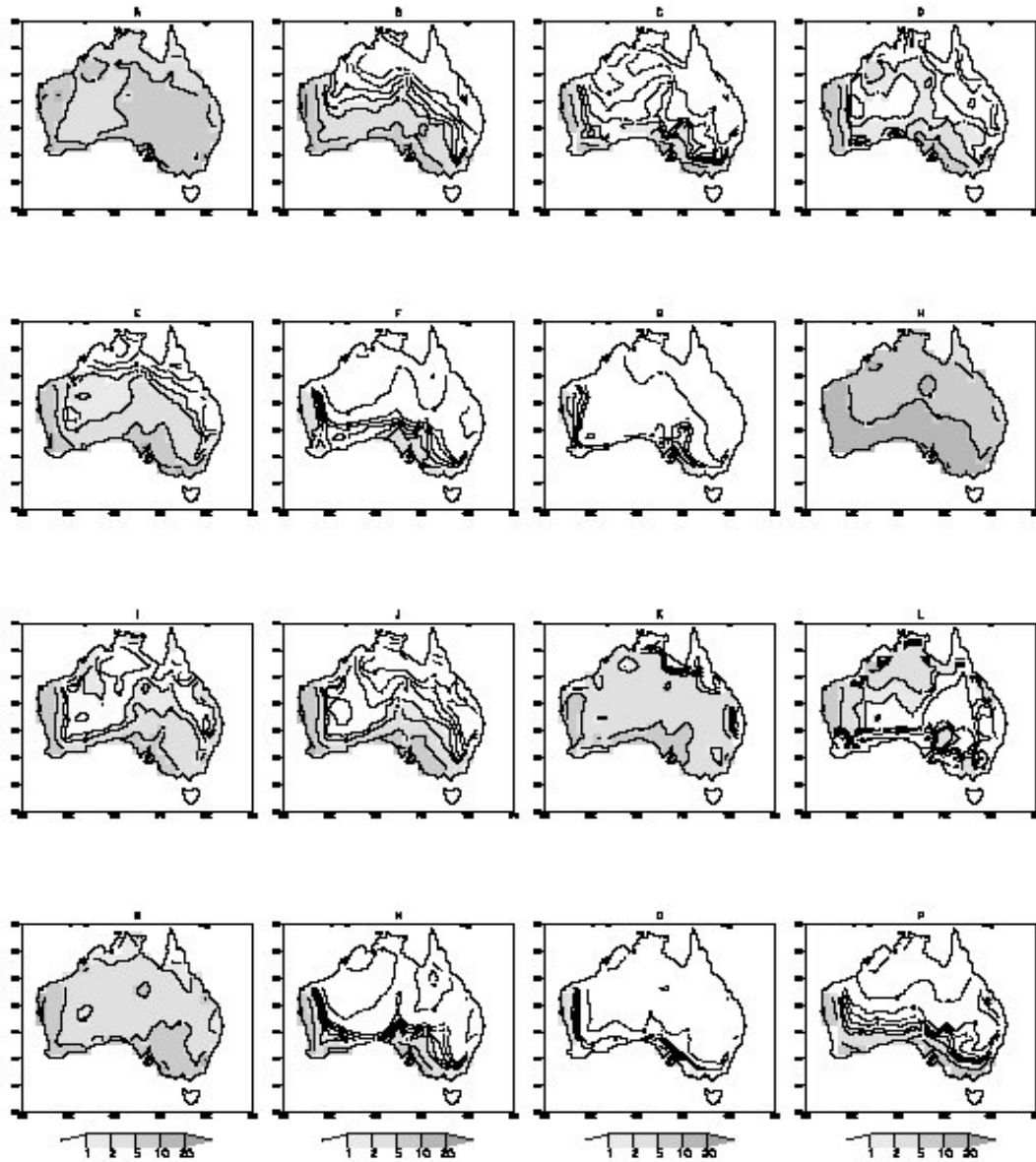
Figure 5(a): Tmax bias in DJF. Unit: K.

5(b) JJA
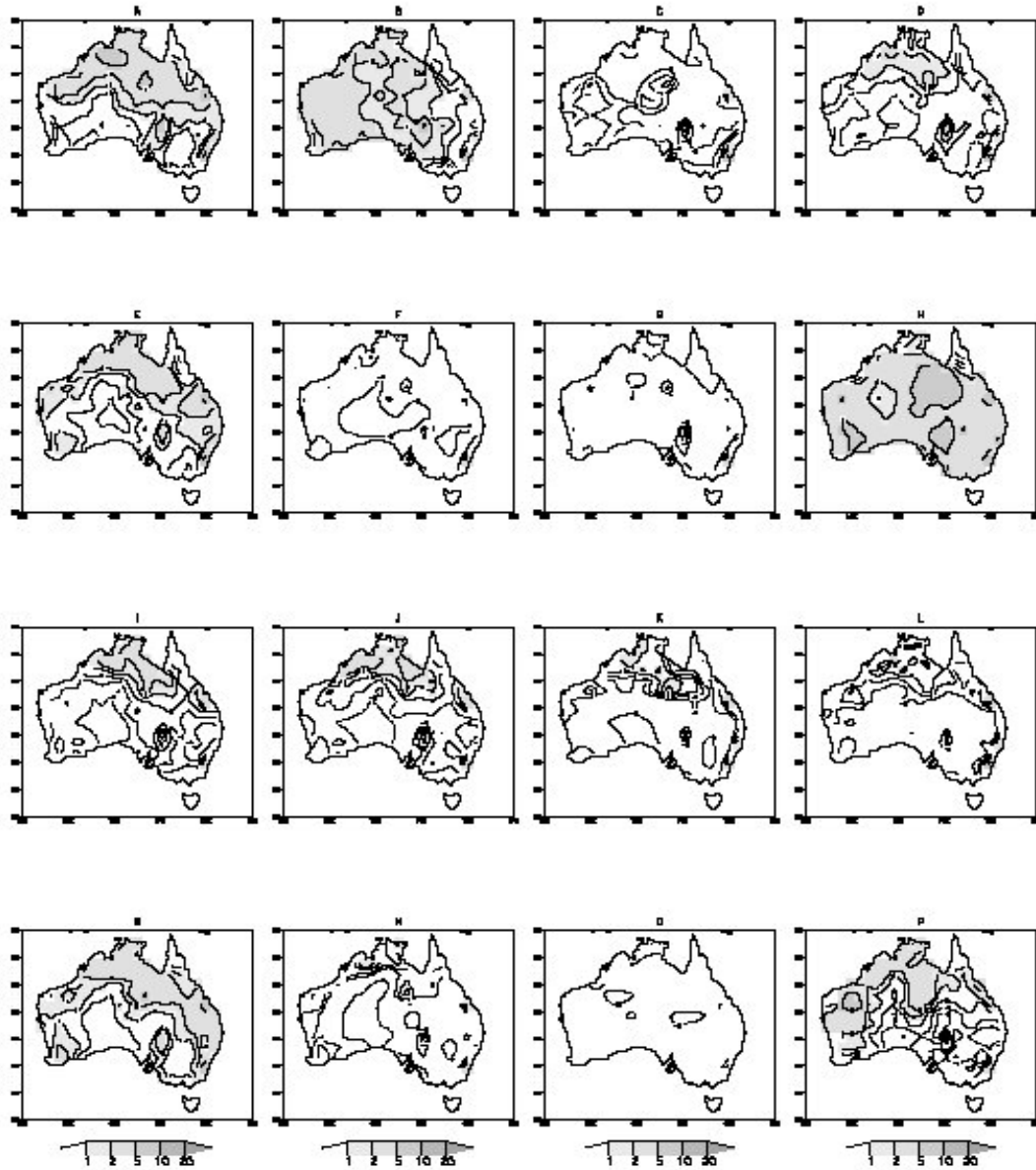
Figure 5(b): Tmax bias in JJA. Unit: K.

Figure 6: As Figure 3 but for monthly mean daily maximum surface temperature (K).
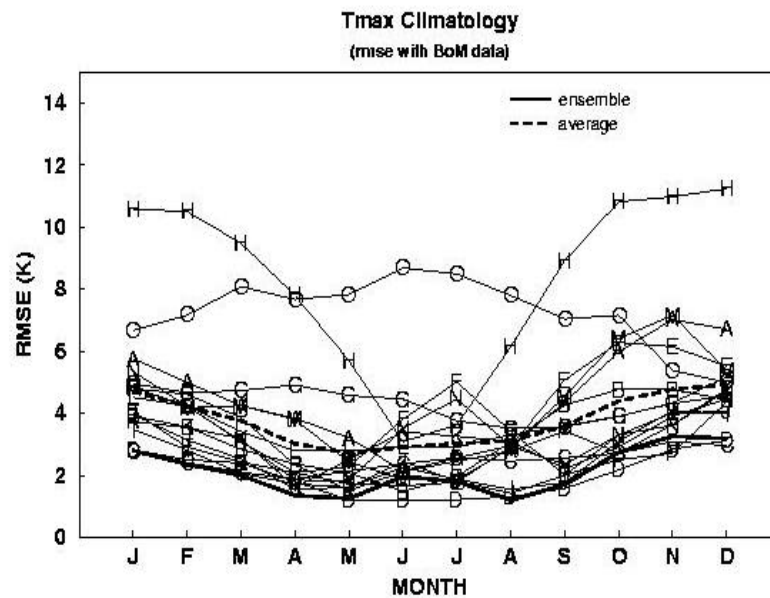
Figure 6: Root–mean–square–error (rmse) of Tmax climatology against the BoM observations (K).

Figure 7: As Figure 4 but for monthly mean daily maximum surface temperature (K).
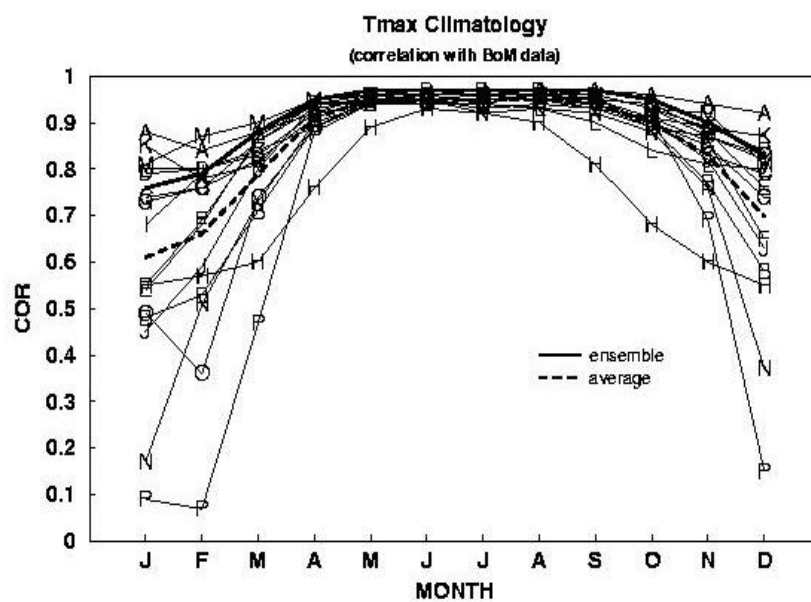
Figure 7: Spatial correlations of Tmax climatology against the BoM observations.

Figure 8: As Figure 2 but for monthly mean daily minimum surface temperature (K).

(a) DJF;
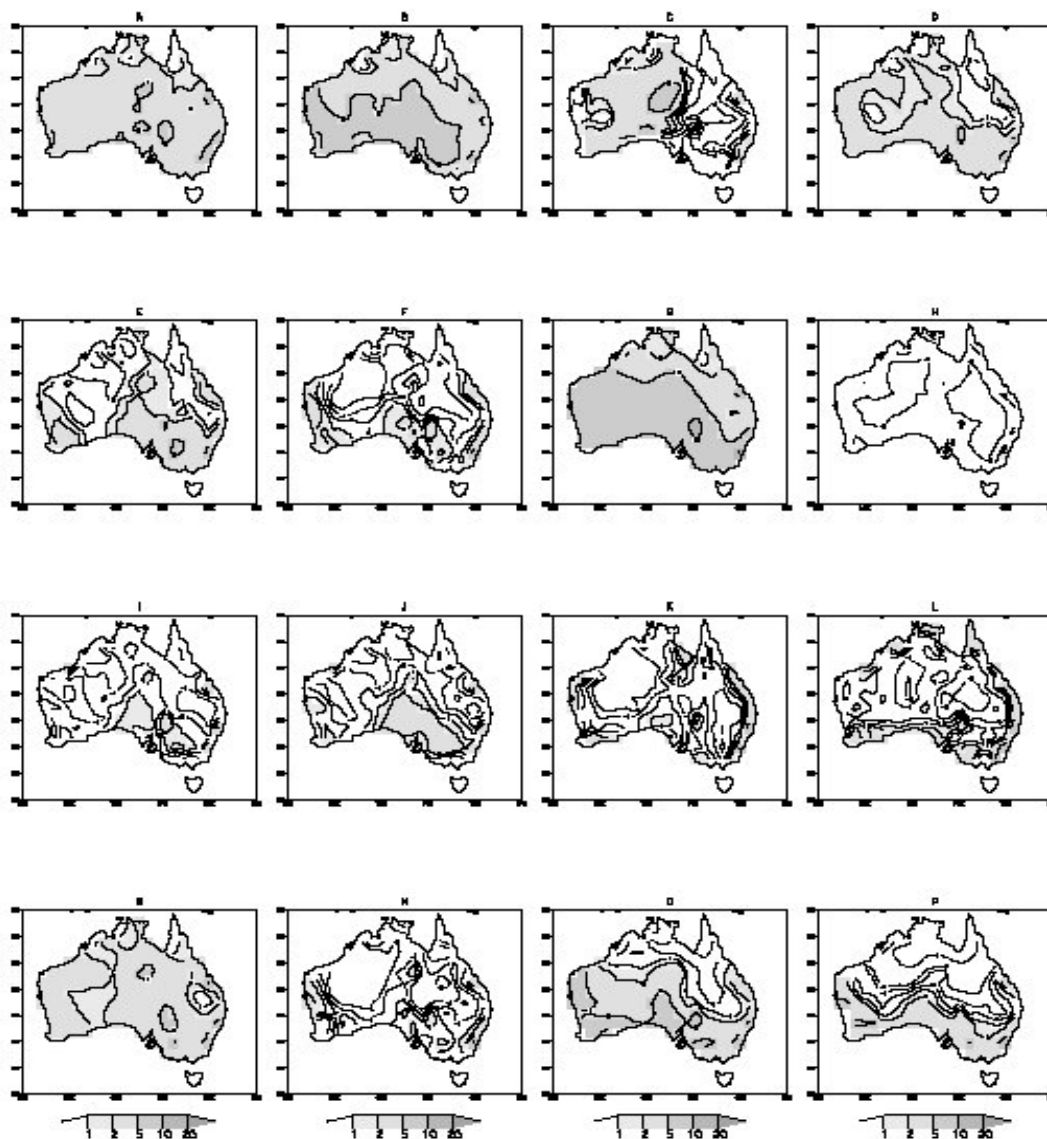
Figure 8(a): Tmin bias in DJF. Unit: K.

8(b) JJA.
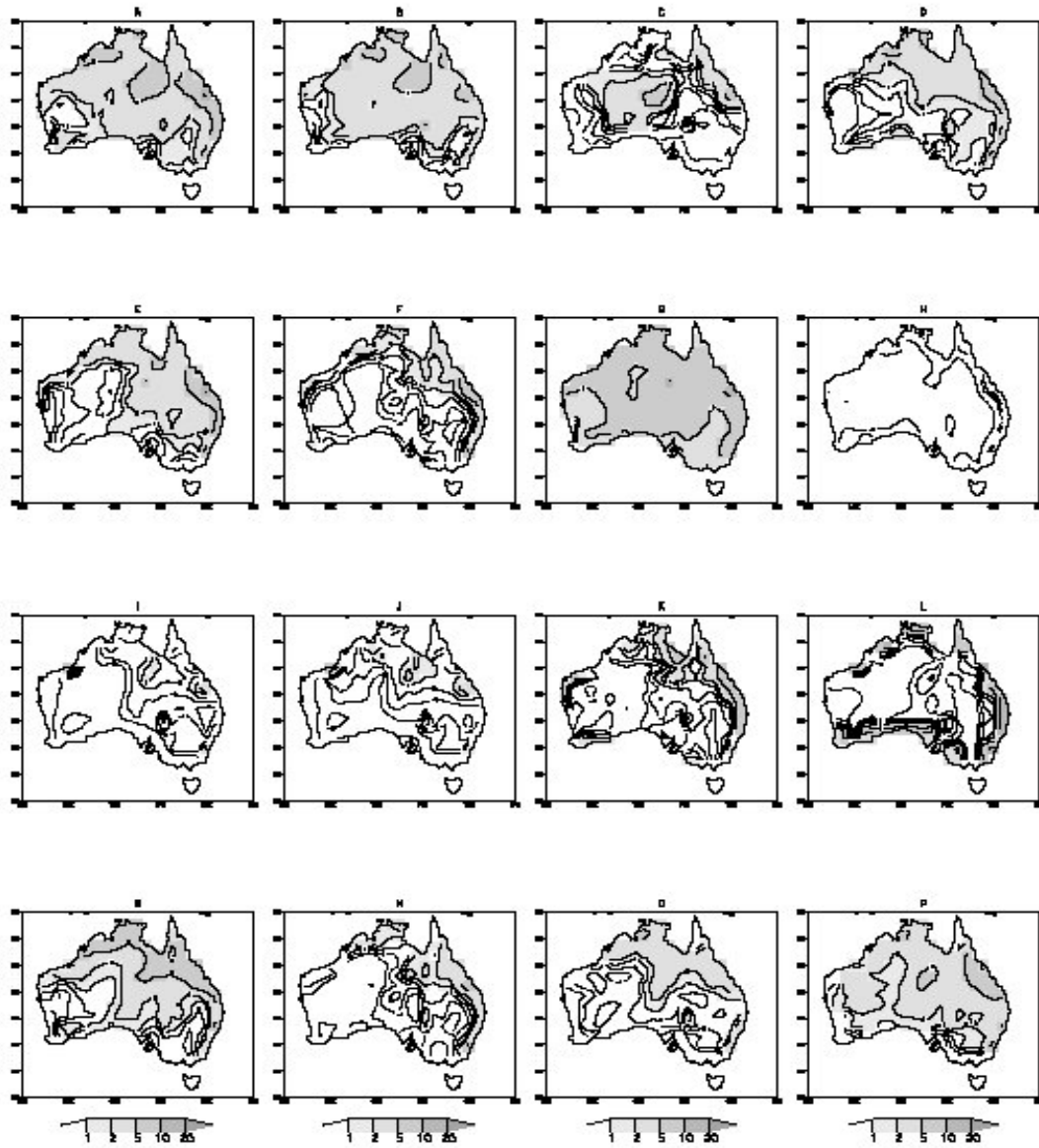
Figure 8(b): Tmin Bias in JJA. Unit: K.

Figure 9: As Figure 3 but for monthly mean daily minimum surface temperature (K).
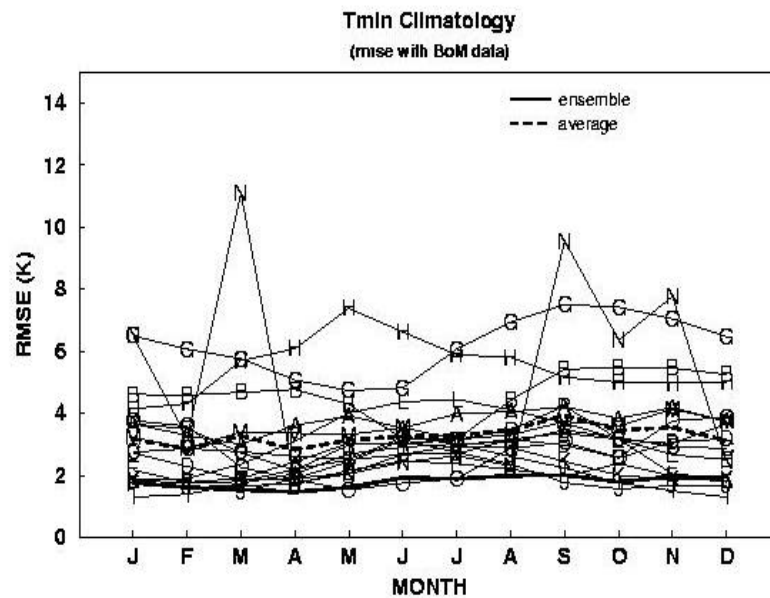
Figure 9: Root—mean—square—error (rmse) of Tmin climatology against the BoM observations (K).

Figure 10: As Figure 4 but for monthly mean daily minimum surface temperature (K).
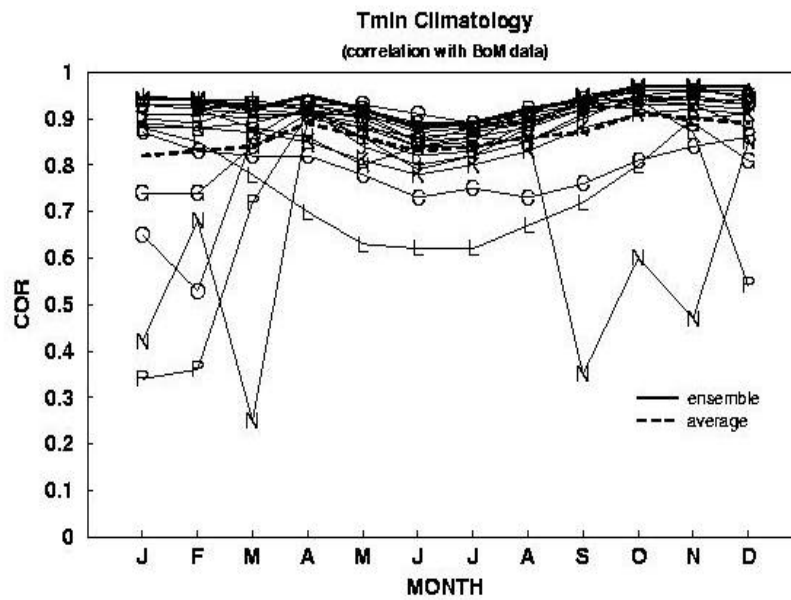
Figure 10: Spatial correlations of Tmin climatology against the BoM observations.

Figure 11: As Figure 2 but for surface temperature daily range as the difference between Tmax and Tmin (K).
(a) DJF;
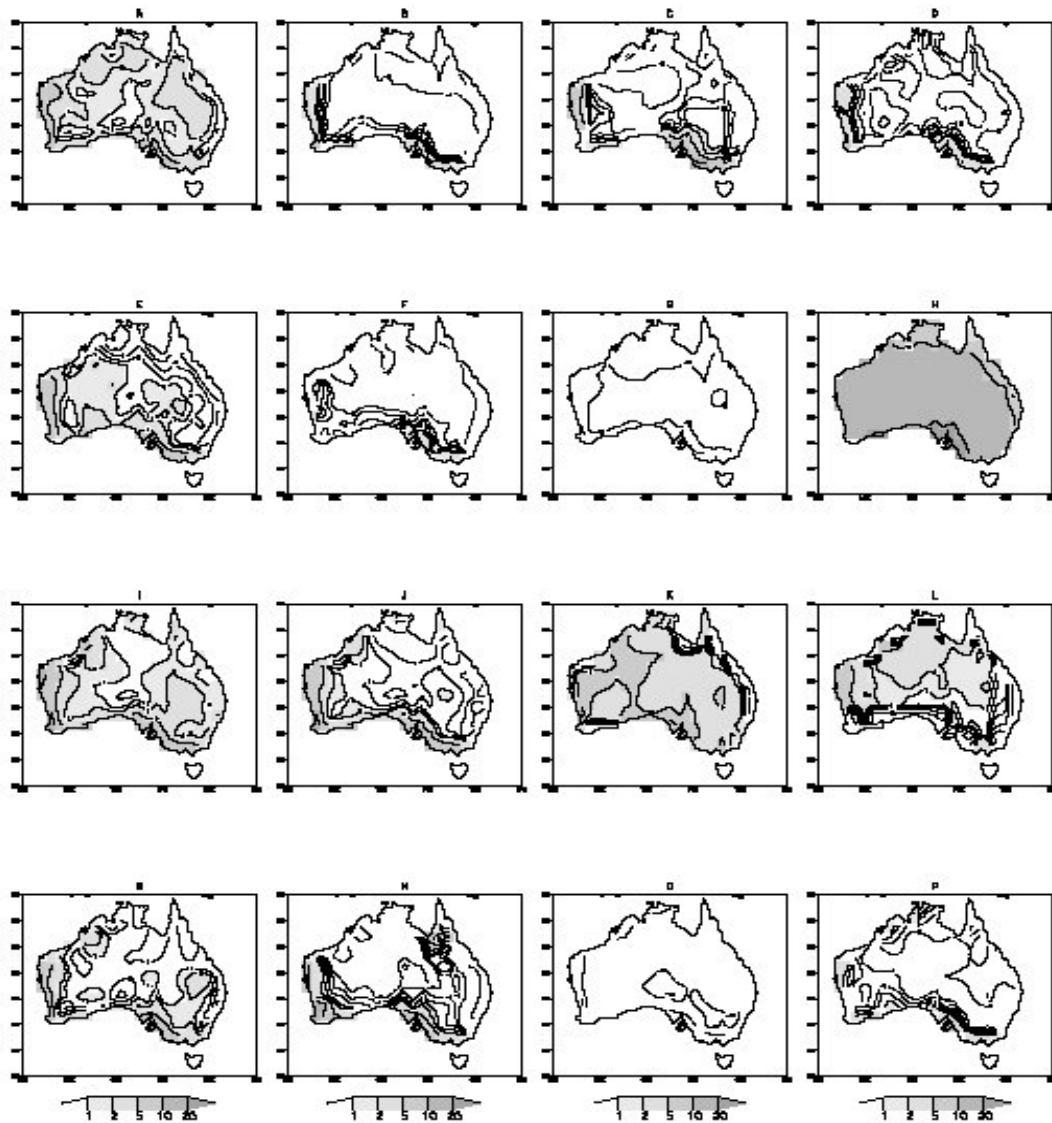
Figure 11(a): Daily temperature range bias in DJF. Units K.

11(b) JJA
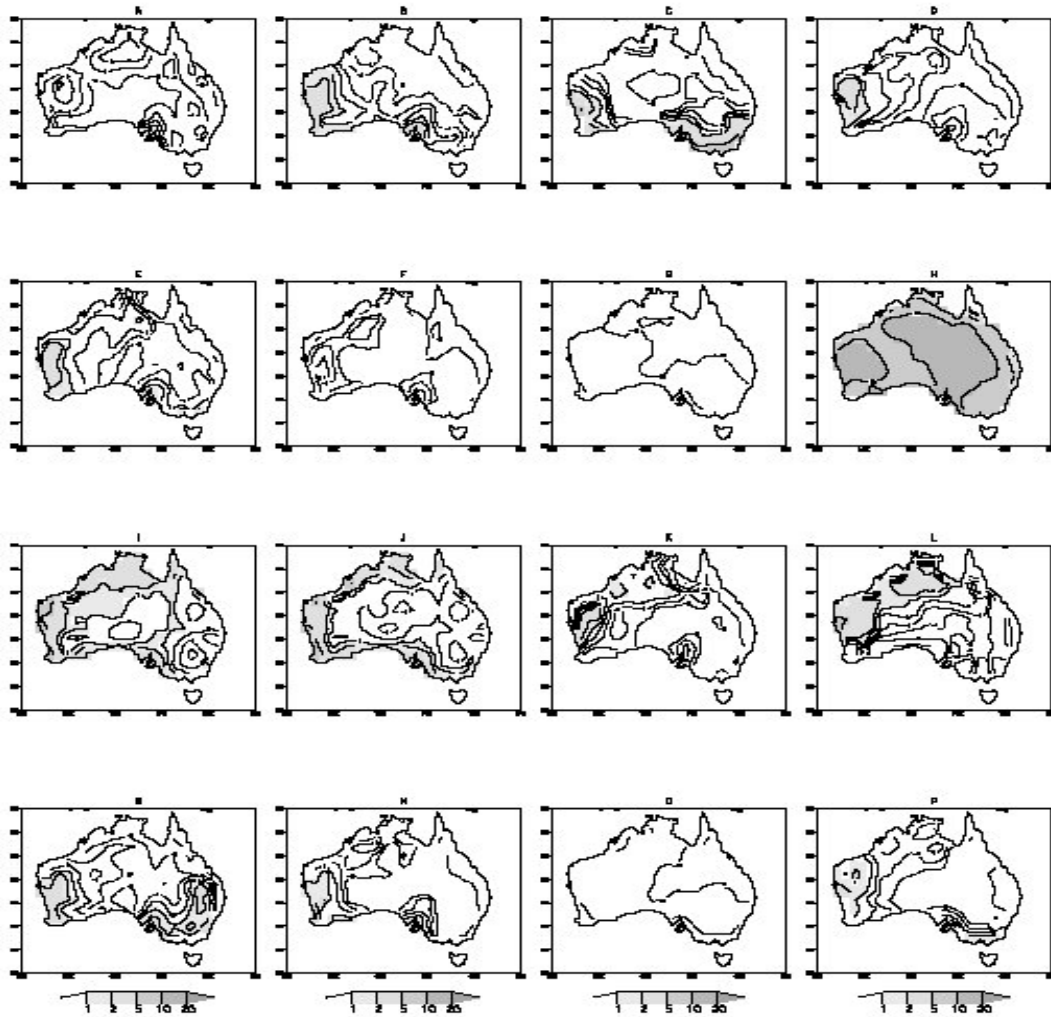
Figure 11(b): Daily temperature range bias in JJA. Unit: K.

.

Figure 12: As Figure 2 but for surface latent heat flux (W m^{-2}).

(a) DJF;
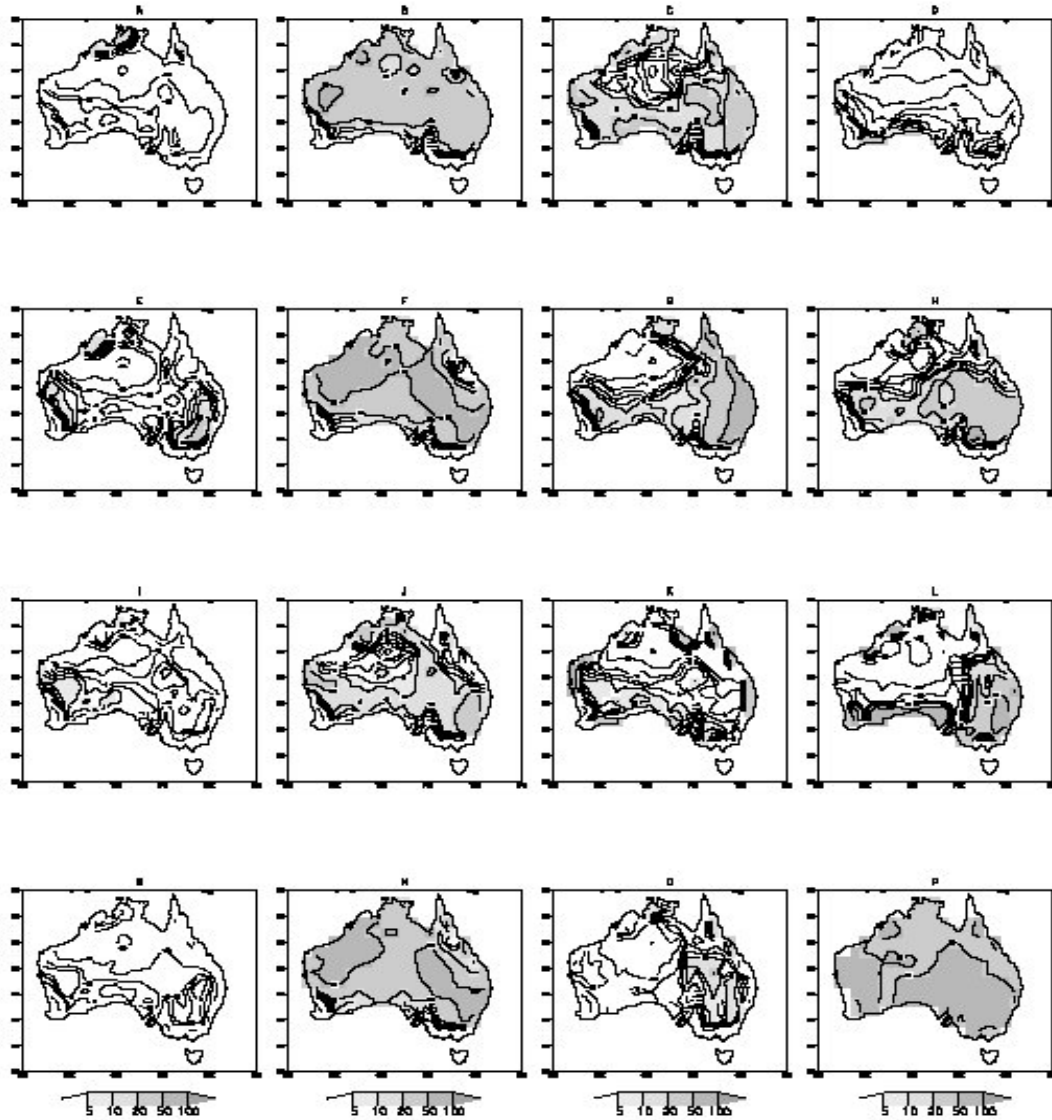
Figure 12(a): Evaporation Bias in DJF. Unit: W/(m**n)

12(b) JJA.
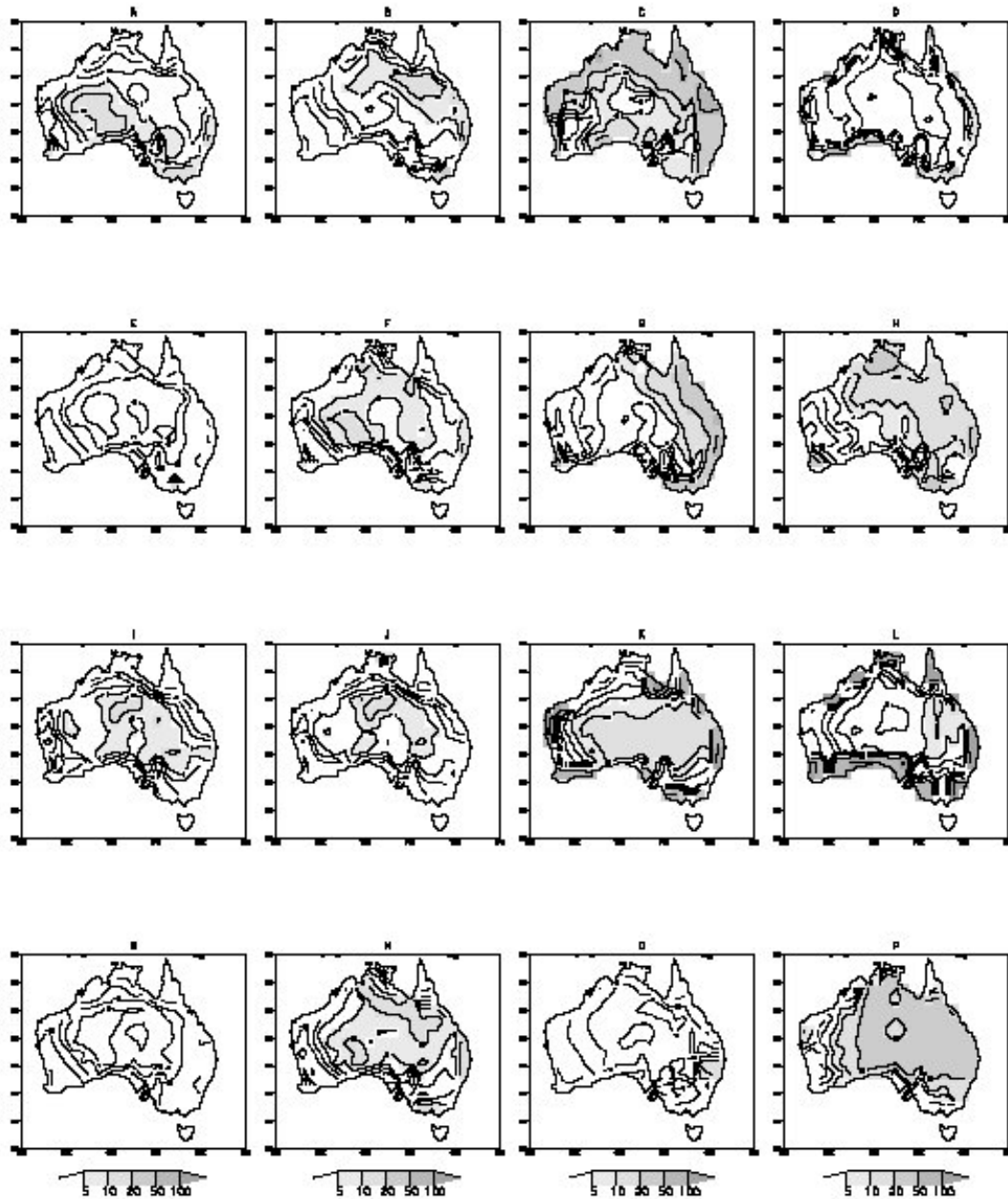
Figure 12(b): Evaporation Bias in JJA. Unit: W/(m*m)

Figure 13: As Figure 3 but for surface latent heat flux (W m^{-2}).
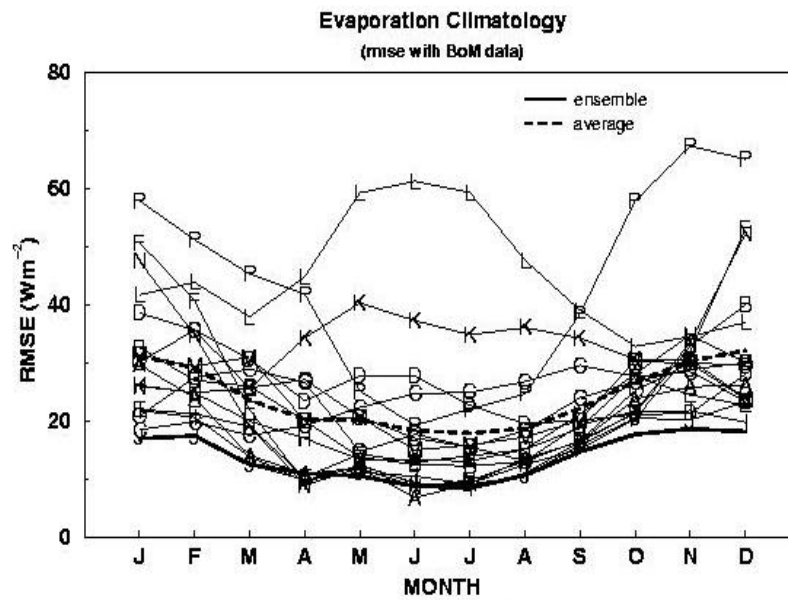
Figure 13: Root–mean–square–error (rmse) of surface evaporation climatology against the BoM's data (W/(m*m))

Figure 14: As Figure 4 but for surface latent heat flux (W m^{-2}).
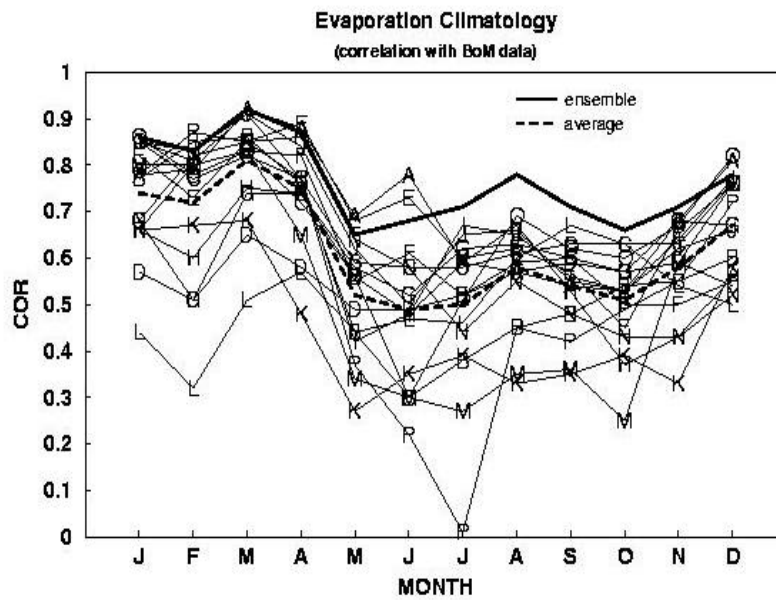
Figure 14: Spatial correlations of surface evaporation climatology against the BoM's data.

Figure 15: LEPS score of 16 AMIP2 models in simulating precipitation anomalies in the 17-yr (1979-1995) period. The BoM observational dataset (1950 to 1999) is used in the calculation. LEPS score in the diagram is divided by 10 with a range of –10 to 10.
(a): DJF;
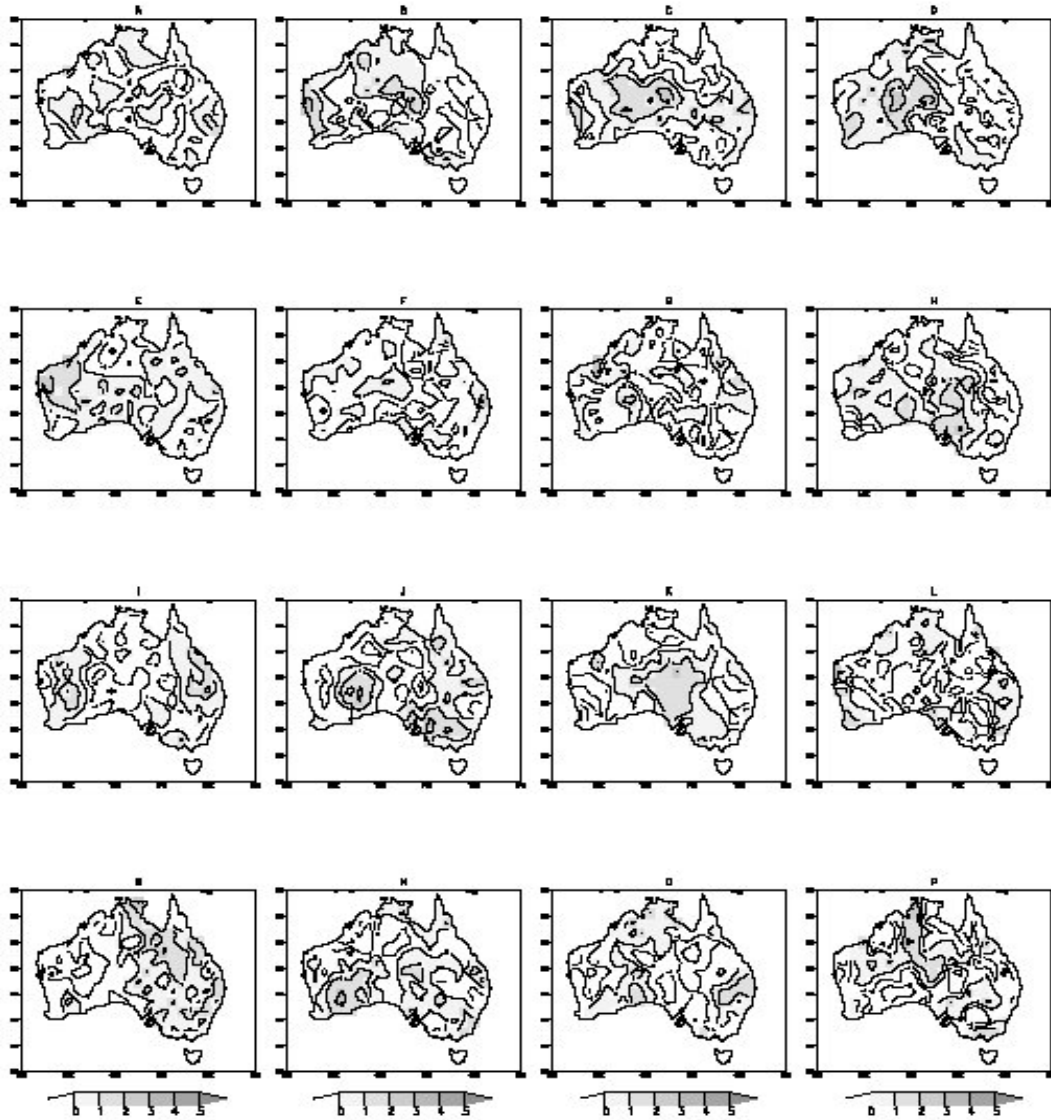
Figure 15(a): Precipitation LEPS scores in DJF.

15(b) JJA.
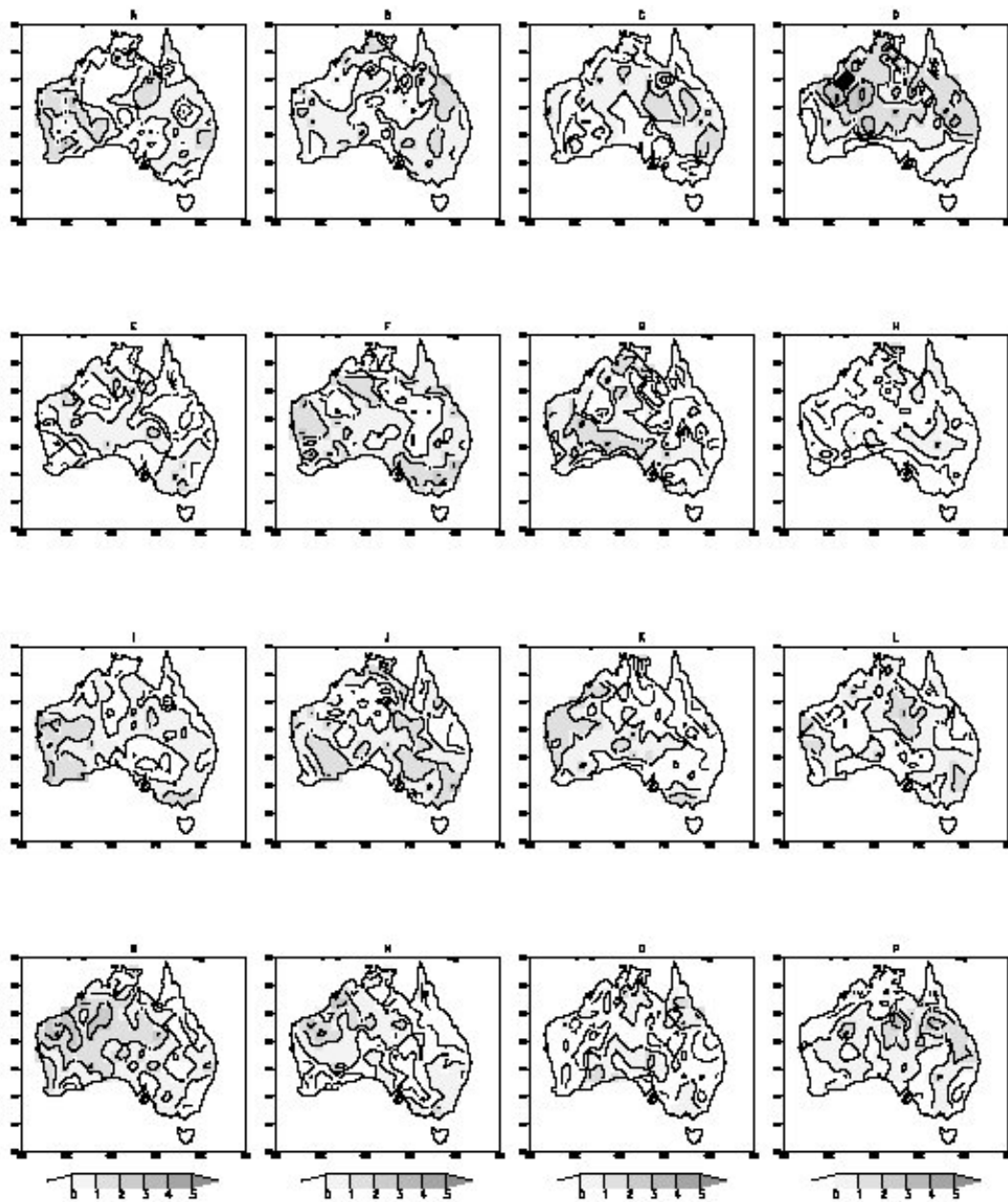
Figure 15(b): Precipitation LEPS scores in JJA.

Figure 16: As Figure 15 but for Tmax.

(a) DJF'
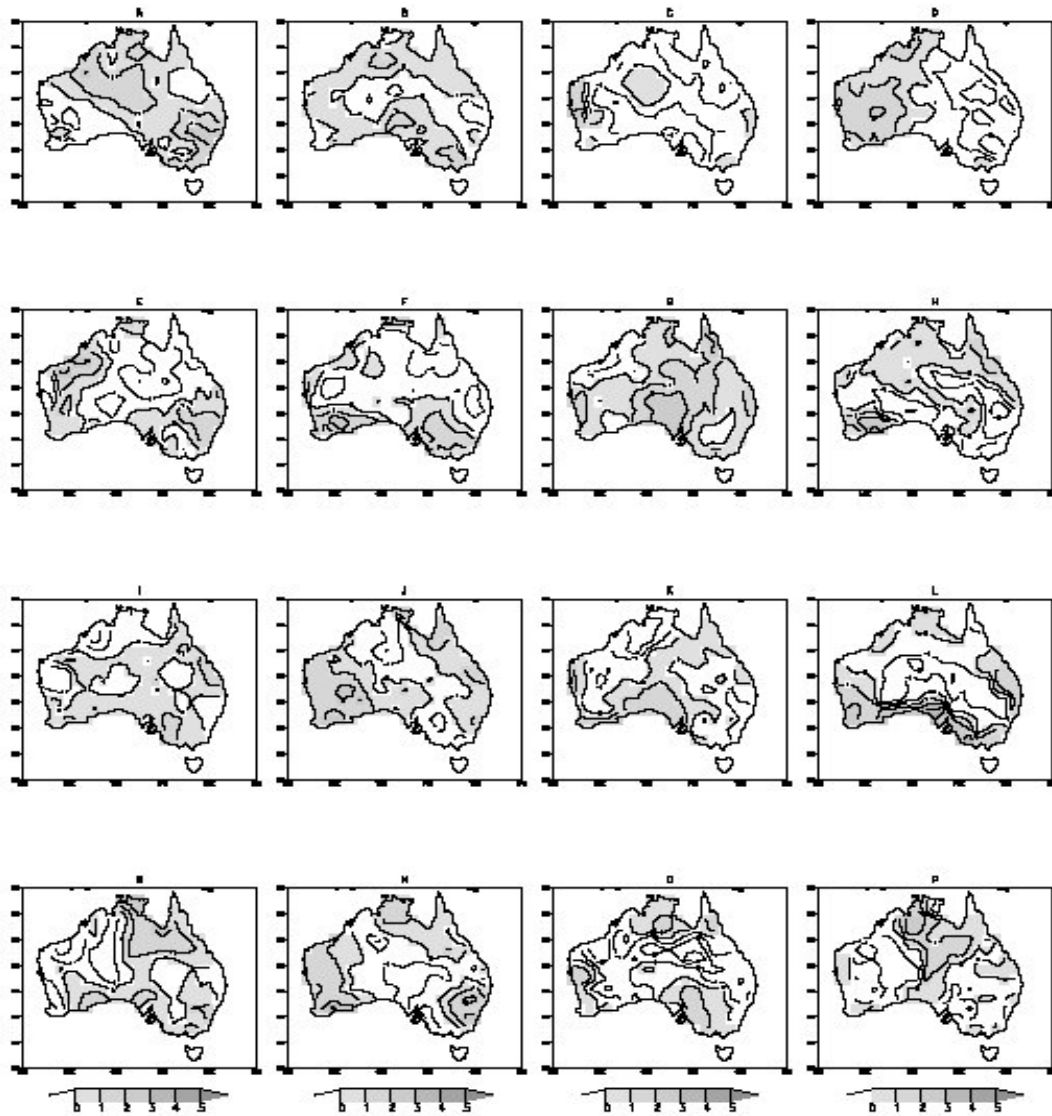
Figure 16(a): Tmax LEPS scores in DJF.

16(b) JJA.
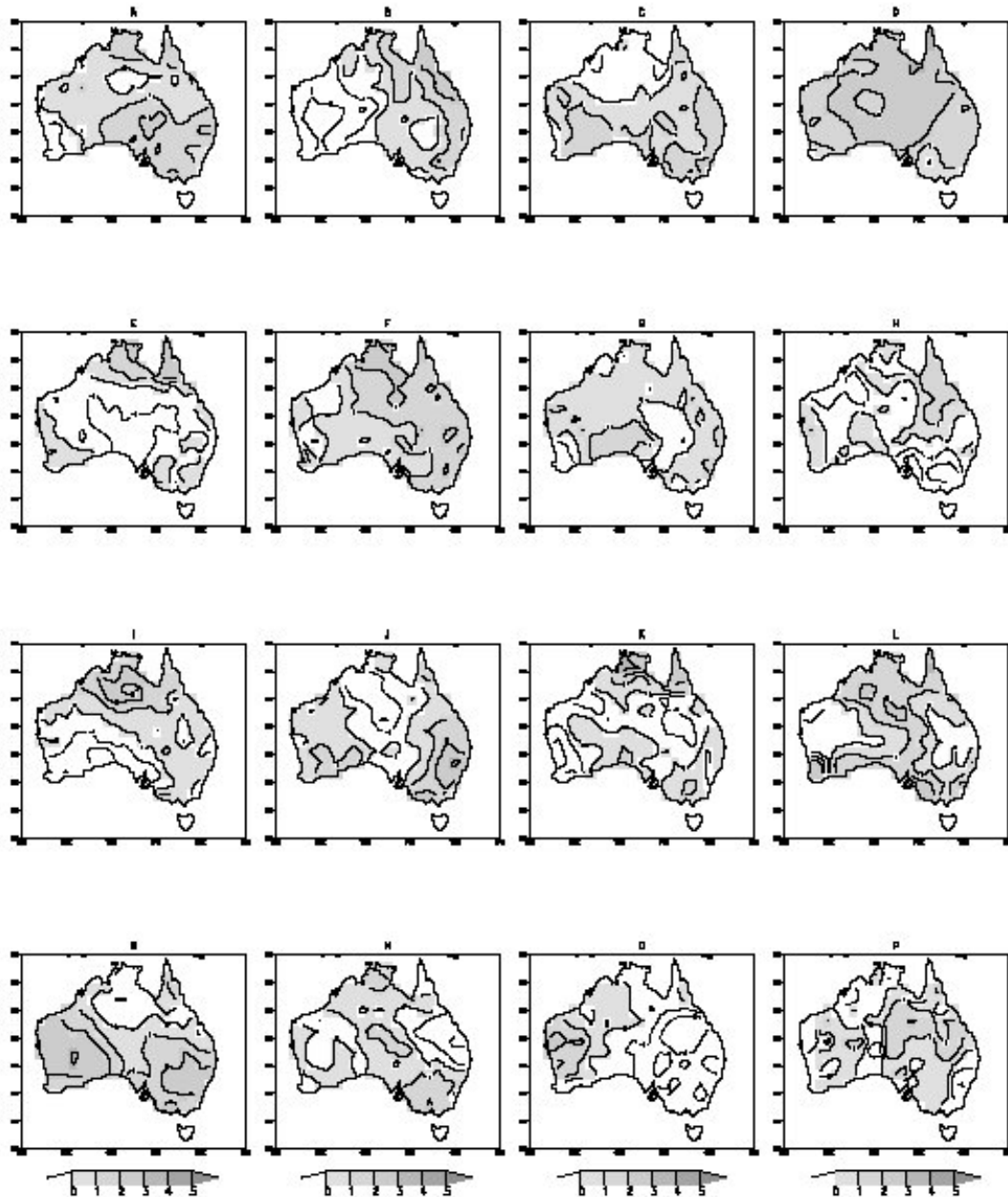
Figure 16(b): Tmax LEPS scores in JJA.

Figure 17: As Figure 15 but for Tmin.

(a) DJF;
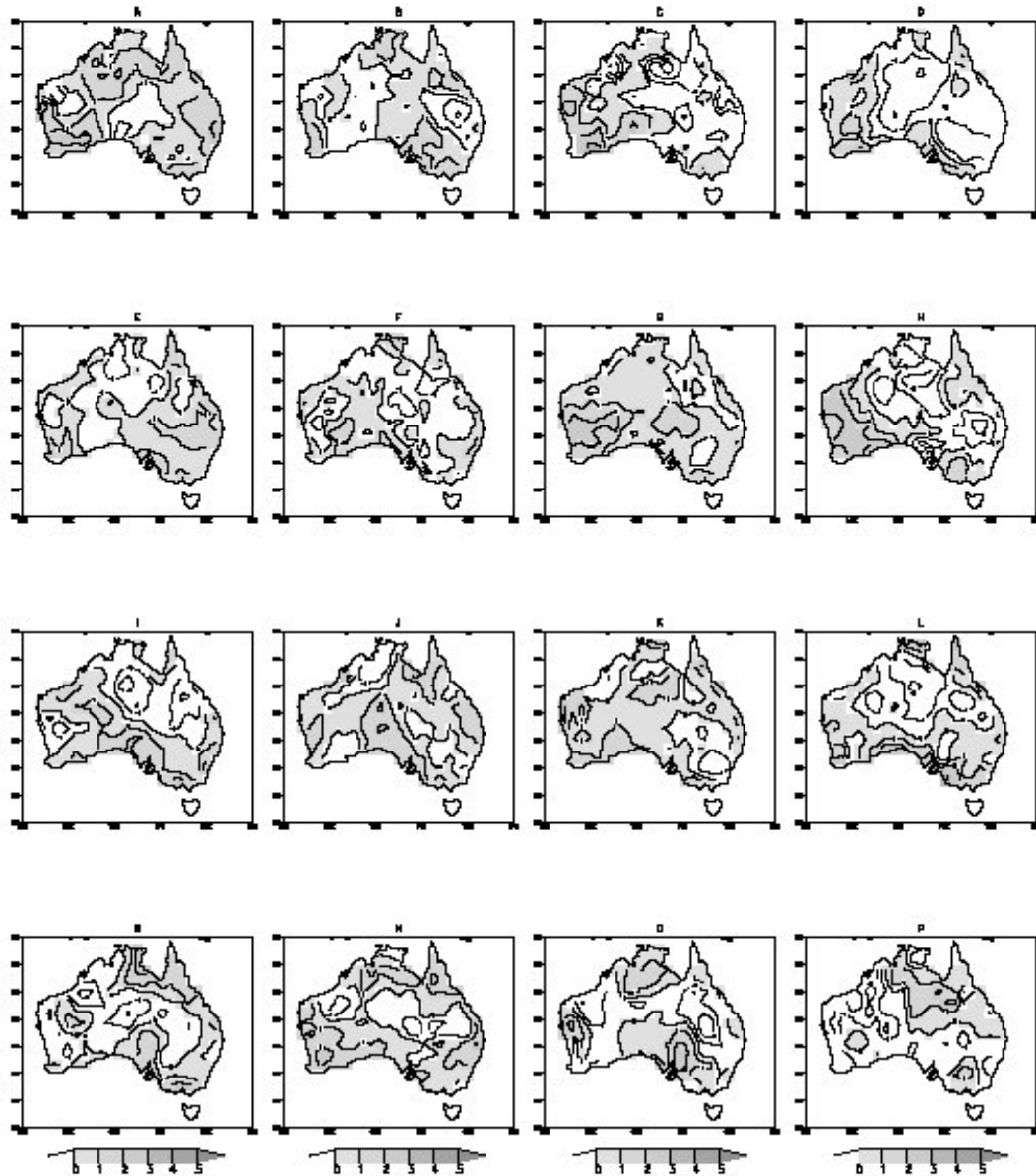
Figure 17(a): Tmin LEPS scores in DJF.

17(b) JJA.

Figure 17(b): Tmin LEPS scores in JJA.

60
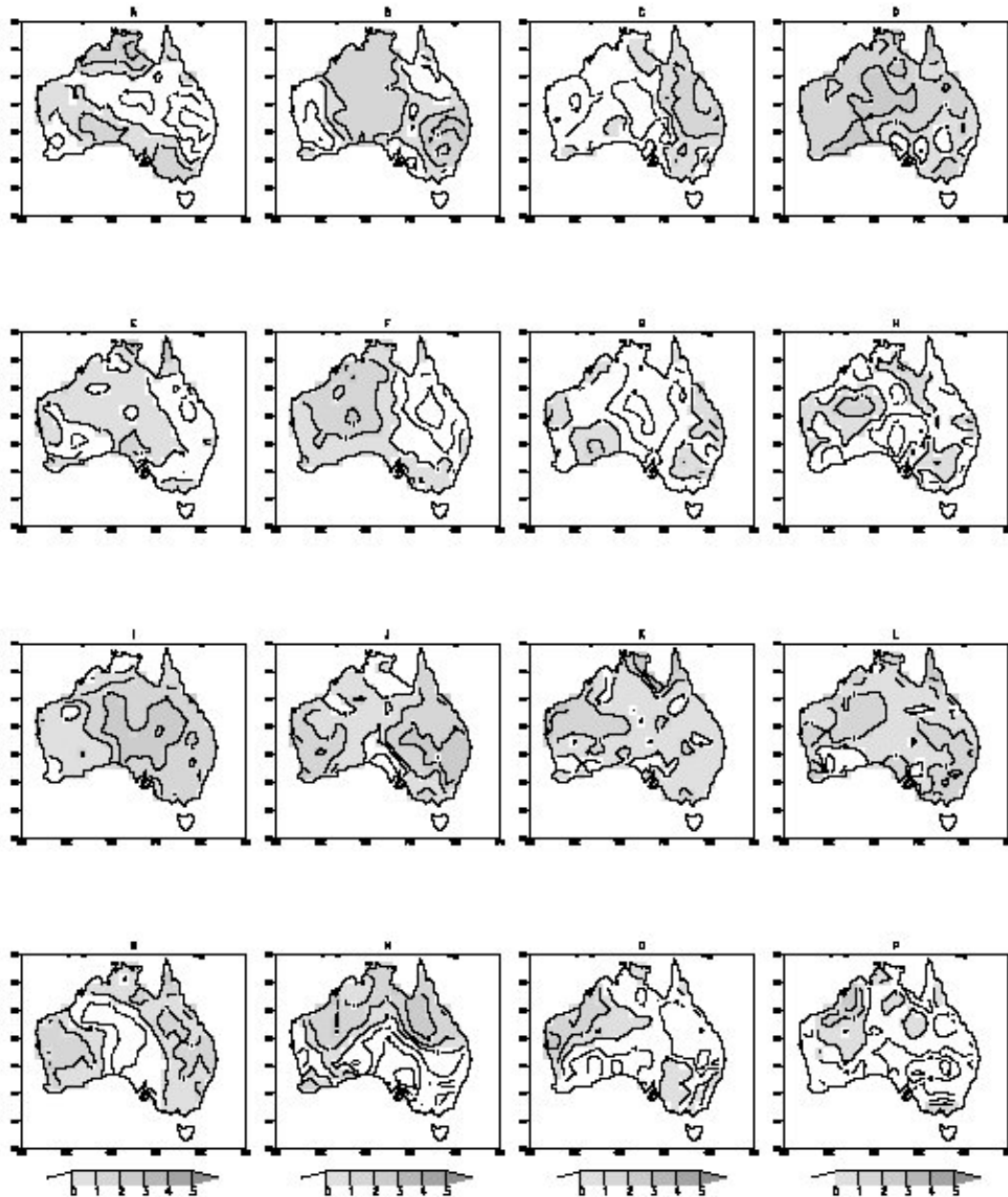
Figure 18: (a): Correlation of each of the 16 models' LEPS scores for the simulation of precipitation anomalies and their scores for the simulation of surface evaporation anomalies over the Australian continent. The heavy dashed line represents the averages of all the 16 models. The heavy solid line represents the results from poor-man ensembles of the 16 models; (b): As (a) but for precipitation and total soil moisture.
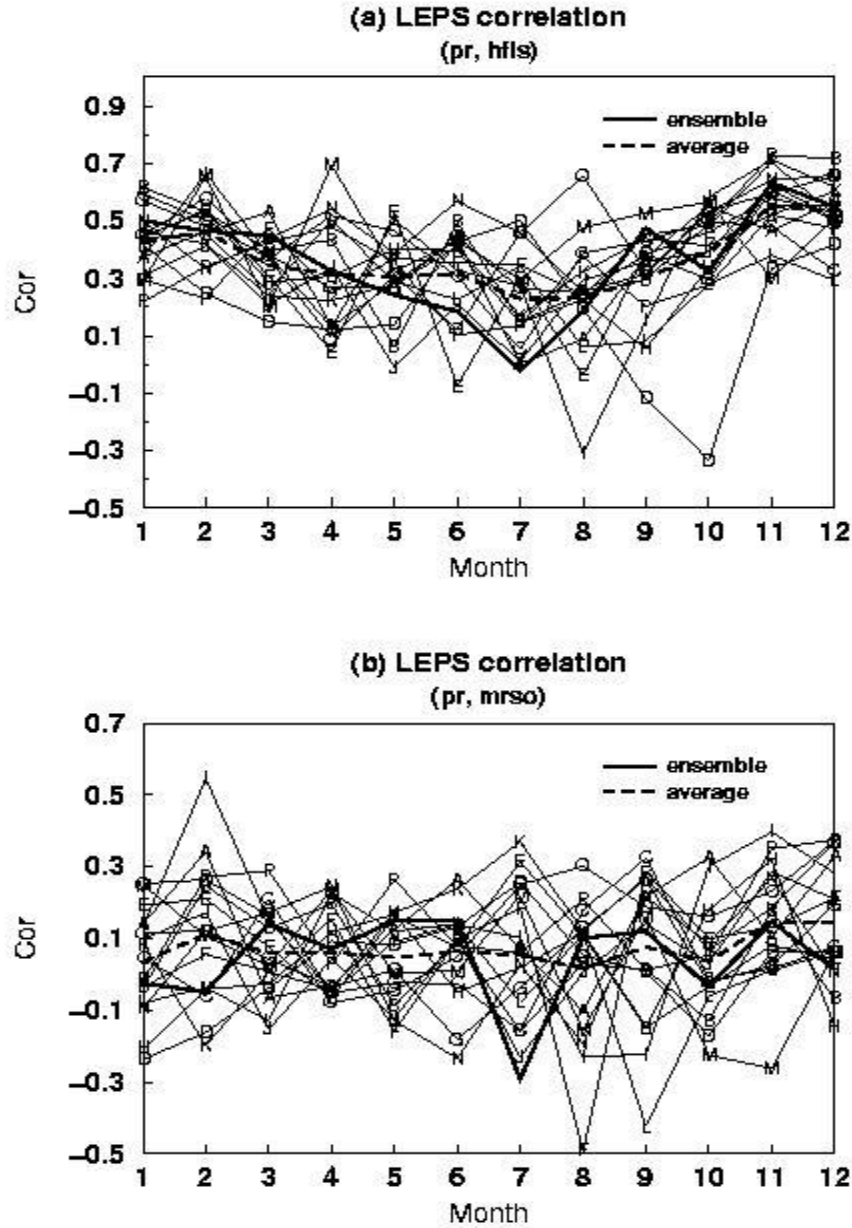
Figure 18: Correlations of LEPS scores from each model over the Australian region. (a): precipitation and surface evaporation; (b) precipitation and soil moisture.

Figure 19: Scatter plots of the correspondence between a single model's scores in simulating precipitation anomalies and surface evaporation anomalies over the Australian region. Each open circle represents results in a single land grid point over the region. (a): DJF; (b): JJA.
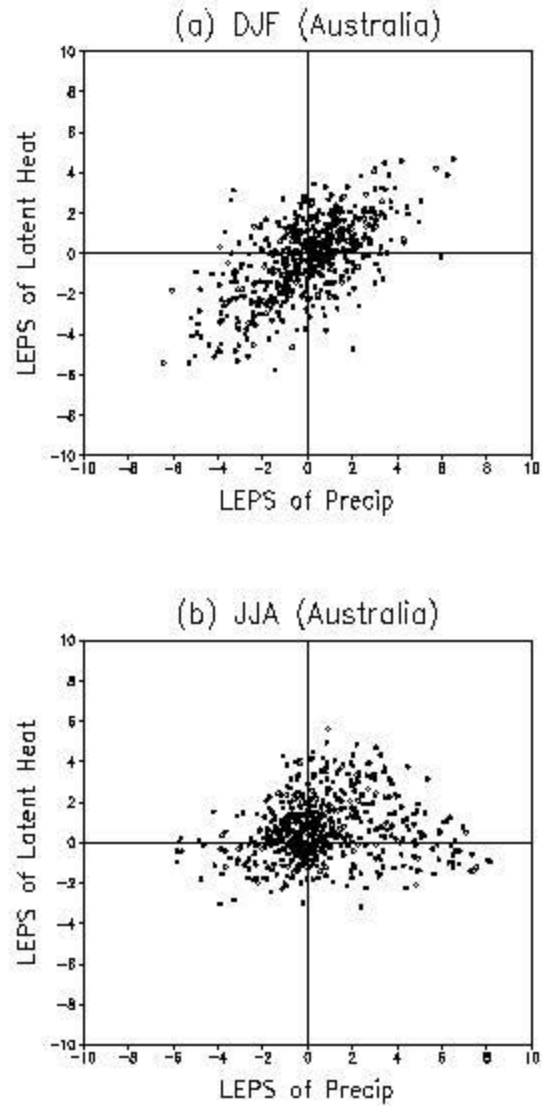
Figure 19: Scatter plots of the correspondence between LEPS scores over the Australian region from a single model.

Figure 20: Three-month lag correlations between total soil moisture anomalies and surface evaporation anomalies, with surface evaporation lagging soil moisture. The seasonal cycle has been removed before the correlation calculations. The value of 0.14 roughly corresponds to a 95% confidence level with 204 samples.
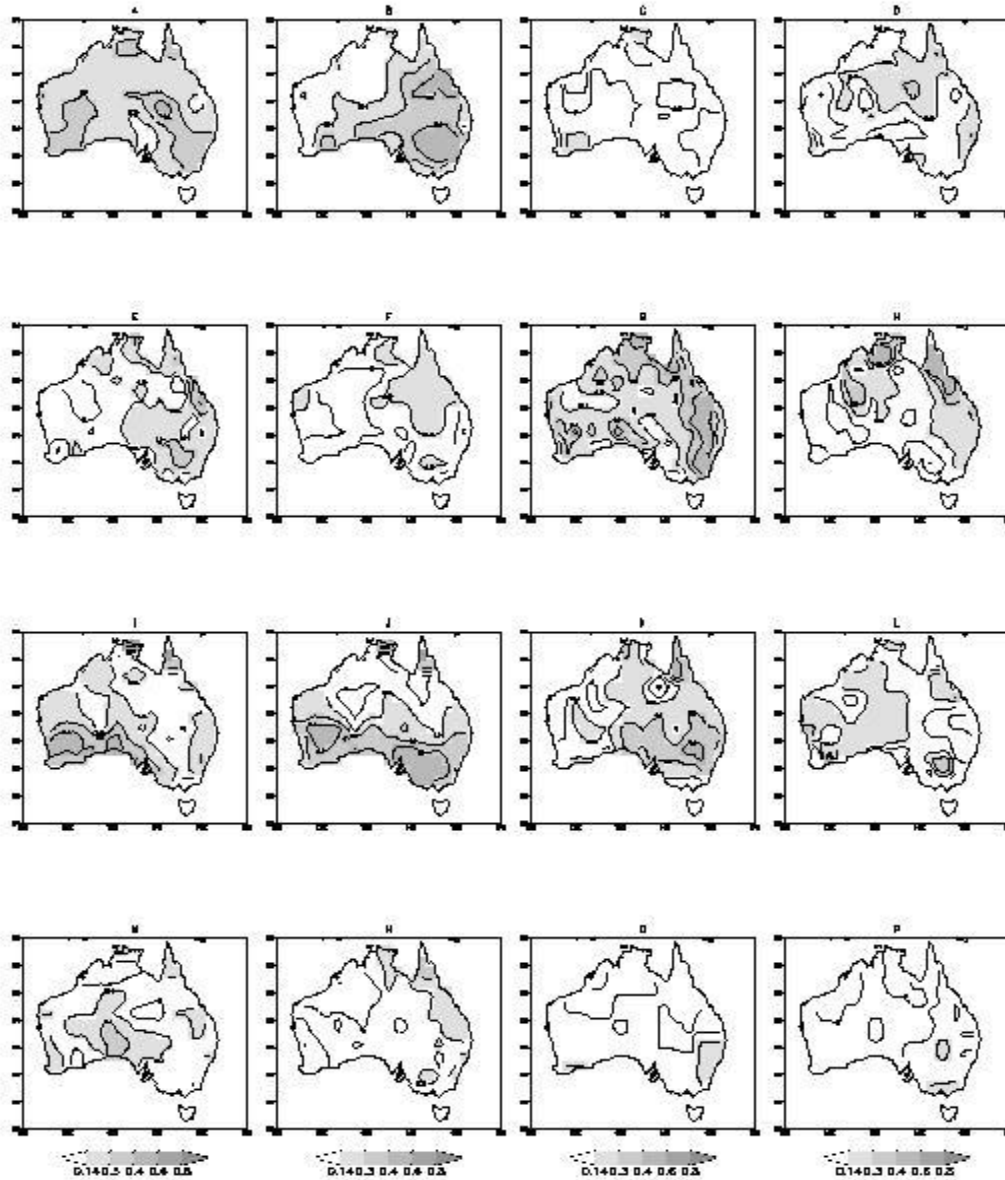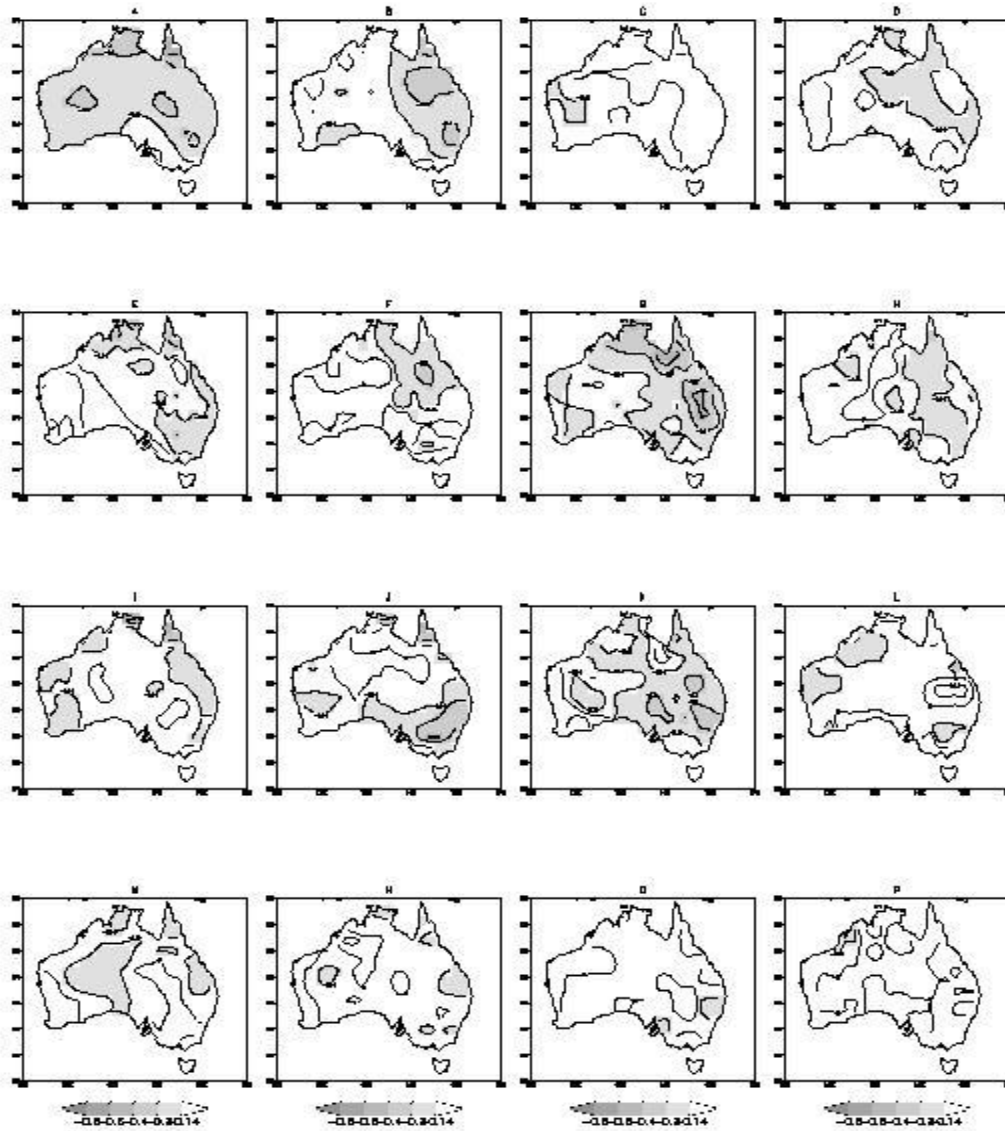

Figure 20: Three—month lag correlation between soil moisture and surface evaporation. Soil moisture leads surface evaporation.

Figure 21: As Figure 20 but for the correlation between total soil moisture and surface air temperature.

Figure 21: Three—month lag correlation between soil moisture and surface temperature.
Soil moisture leads surface temperature.

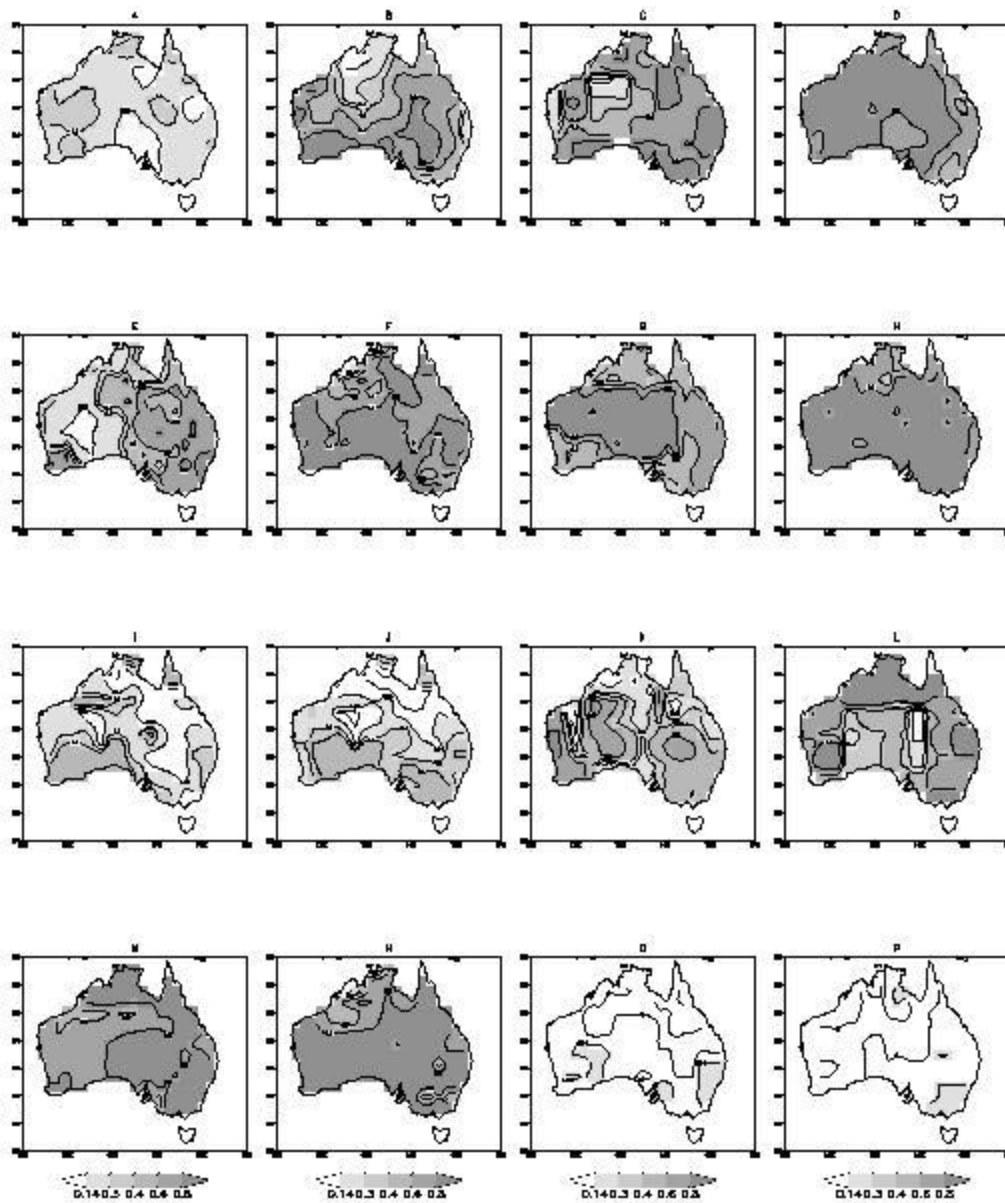Figure 22: As Figure 20 but for the auto-correlation of total soil moisture anomalies.

Figure 22: Three-month auto correlation of soil moisture.

Figure 23: Variation of soil moisture anomalies after removing the seasonal cycle simulated in three models over the location 25 ^{o}S and 135 ^{o}E.
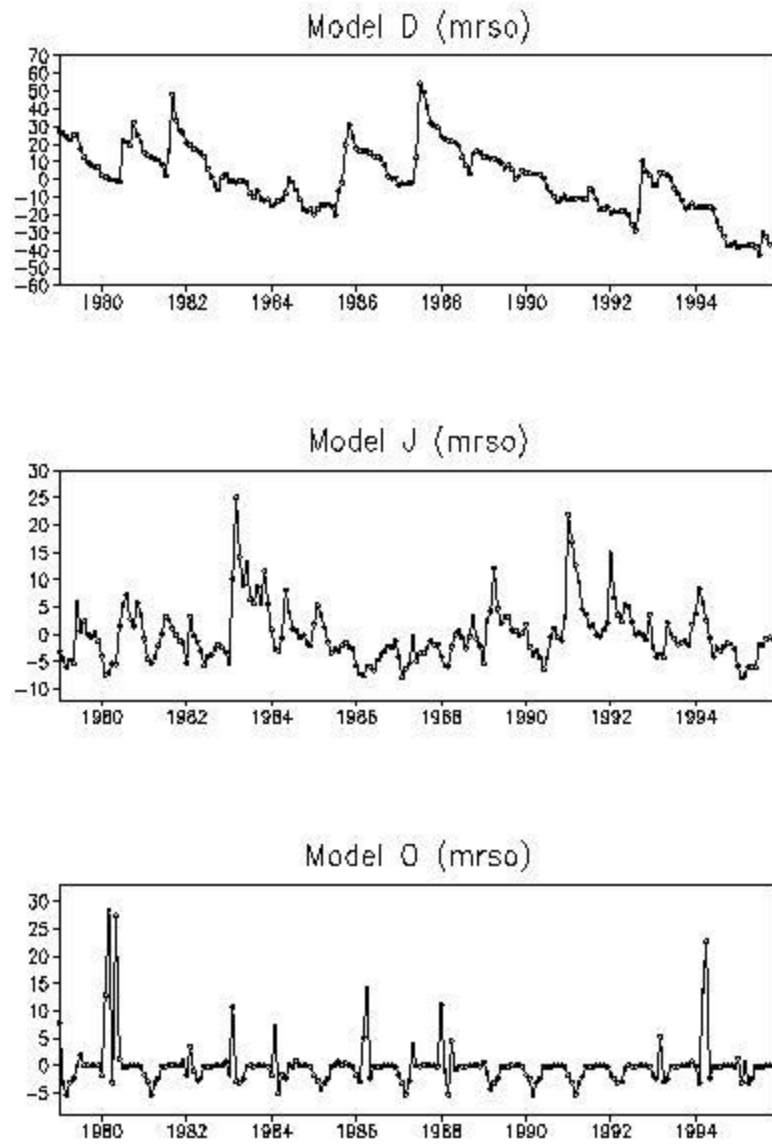

Figure 23 Soil moisture variations in three models over location of 25S and 135E.

Figure 24: (a): Area-averaged lag correlations between total soil moisture anomalies and surface evaporation anomalies over the Australian region, with zero to twelve months lag and surface evaporation lagging soil moisture. (b): As (a) but for soil moisture and surface air temperature.
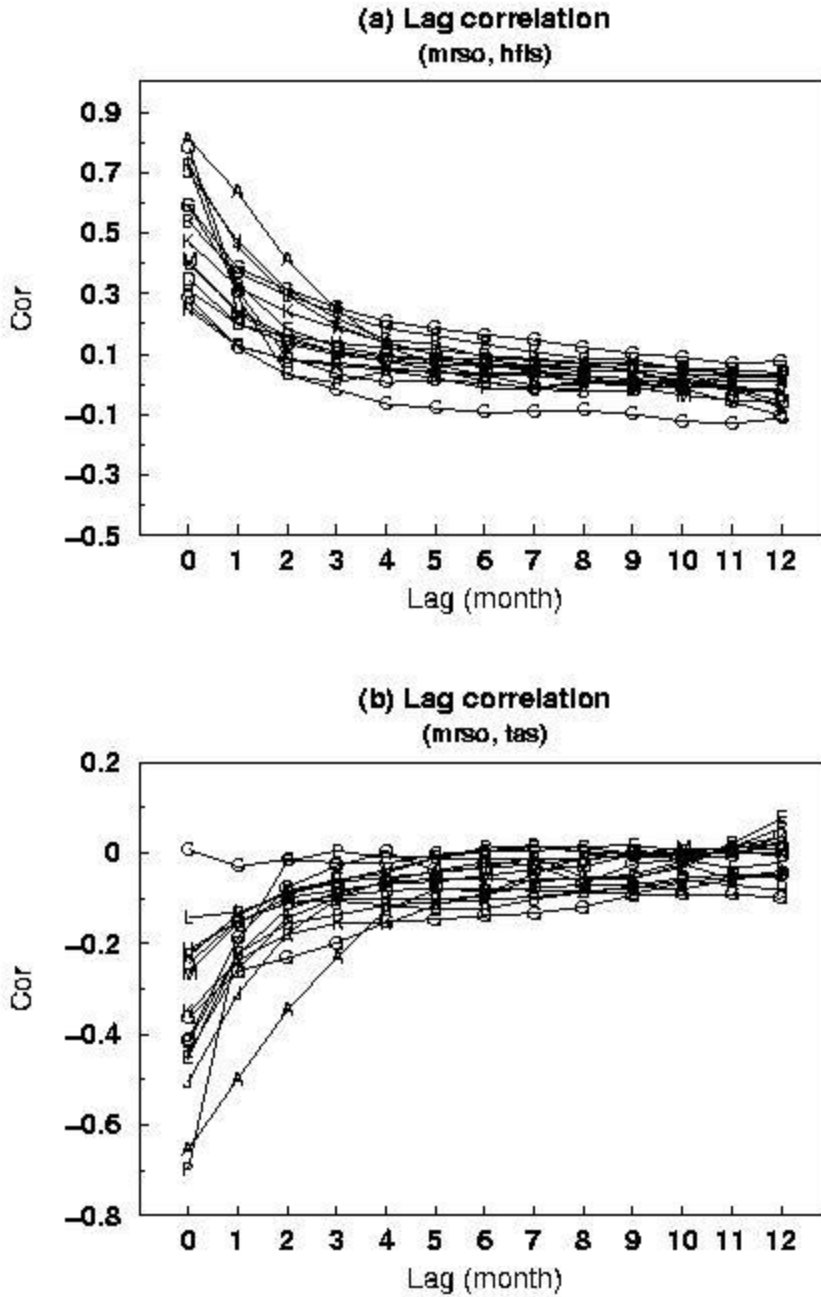

Figure 24: Area—averaged lag correlations between total soil moisture anomalies and surface climate anomalies, over the Australian region.

Figure 25: As Figure 24 but for auto-correlations of total soil moisture (a) and surface air temperature (b).
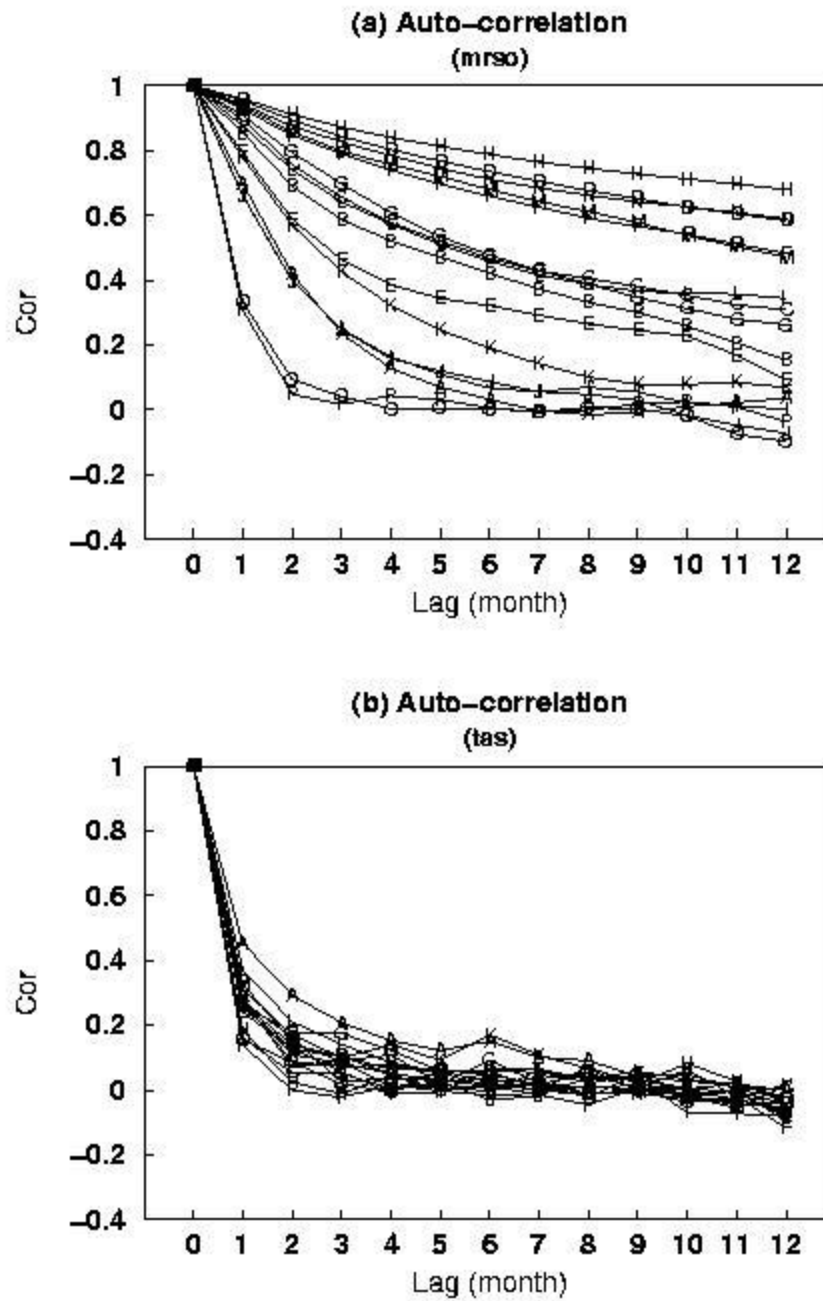

Figure 25: Area-averaged auto correlations of total soil moisture (a) and surface air temperature (b) anomalies over the Australian region.